

ERROR ANALYSIS OF THE FINITE VOLUME ELEMENT
METHOD FOR ELLIPTIC AND PARABOLIC
PARTIAL DIFFERENTIAL EQUATIONS

by

Rick Vance Trujillo

B.S., Metropolitan State College of Denver, 1988

M.S., University of Colorado at Denver, 1991

A thesis submitted to the
University of Colorado at Denver
in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Applied Mathematics

1996

This thesis for the Doctor of Philosophy

degree by

Rick Vance Trujillo

has been approved

by

Thomas F. Russell

Leopoldo P. Franca

Chaoqun Liu

Thomas A. Manteuffel

Stephen F. McCormick

Date_____

Trujillo, Rick Vance (Ph.D., Applied Mathematics)

Error Analysis of the Finite Volume Element Method

for Elliptic and Parabolic Partial Differential Equations

Thesis directed by Professor Thomas F. Russell

ABSTRACT

An *a priori* error analysis of the finite volume element method, a locally conservative, Petrov-Galerkin, finite element method for the numerical solution of elliptic and parabolic partial differential equations arising in fluid dynamics applications, is presented. Existing error estimates apply to discretizations of steady diffusion equations by linear finite elements in two spatial dimensions. These results are extended to steady advection–reaction–diffusion equations and are generalized to polynomial finite elements of arbitrary order in three spatial dimensions and to the full range of admissible regularities for the exact solution. Optimal-order error estimates for h , p , and h - p versions of the method with uniform refinement are derived in a discrete H^1 norm, under minimal regularity assumptions for the exact solution, the finite element triangulation, and the finite volume construction. With additional symmetry assumptions for the finite volumes, multi-dimensional H^1 superconvergence results are obtained for linear finite elements. Using an elliptic projection argument, we outline the analysis of continuous-time and backward Euler and

Crank-Nicolson discrete-time methods for general parabolic equations.

This abstract accurately represents the content of the candidate's thesis. I recommend its publication.

Signed _____
Thomas F. Russell

DEDICATION

To my Family, with love:

To my Father, who finds me when I am lost
and welcomes me home with open arms;

To my elder brother Yeshua, who selflessly pours out his lifeblood
to save me and bears my burdens upon his shoulders;

To my dear Comforter, who always illuminates the way of
truth and life in the darkest nights;

To my Mother, who graciously intercedes on my behalf
and keeps me close to her heart;

To my Guardian, who subtly guides and protects me
all the days of my life;

To my Brothers and Sisters, who encourage and admonish me
with the example of their lives.

Please never let me be separated from thee,
my precious loving Family!

CONTENTS

<u>Chapter</u>	
1.	The Finite Volume Element Method 1
1.1	Introduction 1
1.1.1	Contents of the Thesis 3
1.2	Background 7
1.3	Numerical Methods 9
1.4	The Finite Volume Method (FV) 12
1.5	The Finite Element Method (FE) 14
1.6	The Finite Volume Element Method (FVE) 18
1.7	Historical Development 22
1.8	Summary 25
2.	Theoretical Preliminaries 26
2.1	Triangulation 26
2.1.1	FE Triangulation 27
2.1.2	FVE Primal Mesh 28
2.2	Volumization 32
2.2.1	Volume Construction 32
2.2.2	Volume Symmetry 38
2.2.3	Volume Regularity 38
2.3	Versions h , p , and h - p 41
2.4	FVE vs. FE 43

2.5	Sobolev Spaces	44
2.5.1	Integer-Order Spaces	45
2.5.2	Fractional-Order Spaces	46
2.5.3	Sobolev Embedding	46
2.5.4	Trace Theorem	47
2.5.5	Poincaré Inequality	47
2.6	Finite Element Analysis	49
2.6.1	Boundedness	50
2.6.2	Approximation Theory	51
2.6.3	Ellipticity	52
2.6.4	Error Analysis	52
2.6.4.1	Error Splitting	52
2.6.4.2	Zero Property	53
2.6.4.3	Error Estimate	53
2.7	Summary	54
3.	Analysis for Diffusion Equations	55
3.1	Preliminaries	56
3.2	Motivation	59
3.3	Discrete Norms	61
3.3.1	Discrete L^2 Norm	62
3.3.2	Discrete H^1 Norm	63
3.3.3	Discrete Poincaré Inequality	64
3.4	Boundedness	65
3.5	Bramble-Hilbert Lemma	68

3.5.1	Applications	69
3.6	Approximation Theory	72
3.6.1	Optimal-order Approximation	72
3.6.2	Superconvergent Approximation	77
3.6.2.1	Extension to Three Dimensions	81
3.6.2.2	Degenerate Volume Paradigm	84
3.7	Ellipticity	87
3.8	Error Analysis	91
3.8.1	Error Splitting	91
3.8.2	Zero Property	92
3.8.3	Error Estimate	92
3.9	Summary	94
4.	Analysis for General Elliptic Equations	97
4.1	Preliminaries	98
4.2	Boundedness	100
4.2.1	Advection	100
4.2.2	Reaction	101
4.3	Approximation Theory	103
4.3.1	Advection	103
4.3.2	Reaction	104
4.4	Ellipticity	108
4.4.1	Reaction	108
4.4.2	Advection	110
4.5	Error Analysis	113

4.5.1	Reaction-Diffusion	113
4.5.2	Advection-Reaction-Diffusion	114
4.5.3	General Error Estimate	115
4.6	Elliptic Projection	116
4.7	Summary	118
5.	Analysis for Parabolic Equations	126
5.1	Numerical Methods	128
5.1.1	Continuous-Time	128
5.1.2	Discrete-Time	129
5.2	Components of Analysis	131
5.2.1	Boundedness	131
5.2.2	Approximation Theory	132
5.2.2.1	Spatial	133
5.2.2.2	Temporal	134
5.2.3	Ellipticity	134
5.2.3.1	Spatial	135
5.2.3.2	Temporal	135
5.3	Error Analysis	137
5.3.1	Error Splitting	137
5.3.2	Zero Property	138
5.3.3	Energy Argument	138
5.4	Summary	141
5.5	Future Work	149
	<u>References</u>	152

FIGURES

Figure		
2.1	An example of the triangulation \mathcal{T}^h	29
2.2	$\mathcal{T}^{h/p}$ restricted to a single element of \mathcal{T}^h	30
2.3	Construction of the sub-volume $v_i(T_p)$	34
2.4	A circumcenter finite volume V_i	36
2.5	A general finite volume V_i	37
2.6	Construction of a pair of elements in $\widetilde{\mathcal{T}^{h/p}}$	40
3.1	Formation of the macro-element $E_{ij} = T_{ij}^1 \cup T_{ij}^2$	79
3.2	An example of $\mathcal{V}^{h/2}$ degenerating to \mathcal{V}^h	85

ACKNOWLEDGMENTS

I gratefully acknowledge the financial support of the CU-Denver Math Department through many teaching assistantships, of Pat Roache through Eco-dynamics Research Associates, Inc. and Sandia National Laboratories of Albuquerque, NM, and of Tom Russell through many research assistantships: most recently through National Science Foundation Grant No. DMS-9312752 and U.S. Army Research Office Grant No. 34593-GS.

I would like to acknowledge some of the people who encouraged and admonished me during research seminar presentations: Teri Barth, Dave Dean, Rick Healy, Randy LeVeque, Tom Manteuffel, Steve McCormick, Suely Oliveira, Jim Otto, Rossen Parashkevov, Pat Roache, Tom Russell, and Hong Wang. I would also like to acknowledge some of the people who encouraged and admonished me in my true vocation as a teacher: Bill Briggs, Tom Kelley, and my students! And a special thanks to Debbie Wangerin, who is always helping behind the scenes.

1. The Finite Volume Element Method

1.1 Introduction

This thesis presents an *a priori* error analysis of the finite volume element method (FVE) for elliptic and parabolic partial differential equations. As introduced by Baliga and Patankar [4] and elaborated by McCormick [39] and in the references cited therein, we can view FVE as a combination of the standard finite volume method (FV), also known as cell-centered finite differences (see Mitchell and Griffiths [36] for details), and the standard Galerkin finite element method (FE) (see Ciarlet [11] for details). In FVE, the governing partial differential equation is posed as a local (integral) conservation law on a set of control or finite volumes—regions of local conservation—that partitions the problem domain, as in FV. The unknown FVE solution and known (boundary and source term) data are systematically discretized by a C^0 piecewise polynomial trial space of degree $p \geq 1$ that is defined on a finite element triangulation with maximum mesh diameter h that partitions the problem domain, as in FE. Thus, the finite volume element method can be viewed as either a systematic generalization of the standard finite volume method or as a locally conservative, as well as globally conservative, (Petrov-Galerkin) variant of the standard finite element method. In analogy to the h , p , and h - p version FE of Babuška (see [2, 3, 27, 51] and the references cited therein), we consider h , p , and h - p version FVE where increased accuracy is obtained by refining the computational mesh parametrized by h and/or by increasing the order p

of the FE polynomial space. Other methods that are either closely related or identical to the h -version FVE include: control volume finite element methods [4], generalized box methods [5], and the method of support operators [48],

Although a large body of theoretical infrastructure and of results for elliptic and parabolic equations exists for both FV and FE, analysis for the h -version FVE is limited to the foundation laid by Cai and McCormick [8], Cai, Mandel, and McCormick [9], and culminating in Cai [10] for the numerical solution of steady diffusion equations by linear finite elements in two spatial dimensions. Starting from this foundation, we extend and generalize FVE error analysis to steady and transient advection–reaction–diffusion equations and to polynomial finite elements of arbitrary order in three spatial dimensions. To construct an error analysis for FVE, we are guided by the large body of FE theory for elliptic and parabolic problems and their solution by arbitrary order finite elements—as contained in Ciarlet [11], Douglas [12], Douglas and Dupont [13, 14], Wheeler [55], and in references cited therein—and by modern FV error analysis for the lowest-order methods based on FE-style arguments for linear finite elements: e.g., Bank and Rose [5], Ewing *et al.* [17, 18], Hackbusch [21], Heinrich [23], Herbin [24], Kreiss *et al.* [31], Lazarov *et al.* [32, 33], Manteuffel and White [35], Morton and Süli [38], Samarskii *et al.* [46], Süli [50], and Weiser and Wheeler [54].

By adapting FV and FE arguments to fit the context of FVE and developing some new arguments unique to FVE, we derive optimal-order error estimates under the full range of admissible regularities for the exact solution for h , p , and h - p version FVE with uniform refinement in a discrete H^1

norm for Lagrange finite elements of arbitrary order—under minimal regularity assumptions for the exact solution, the finite element triangulation, and the finite volume construction. With additional uniformity assumptions for the finite volumes, H^1 superconvergence results are obtained for linear finite elements. Using an elliptic projection argument, we outline the analysis of continuous-time and backward Euler and Crank-Nicolson discrete-time methods for general parabolic equations.

1.1.1 Contents of the Thesis

The remainder of the thesis is organized as follows. In this chapter, we give the necessary introduction for the finite volume element method (FVE) for elliptic equations before we proceed with an analysis of the method. In Section 1.2, we outline how FVE arises from the approximation of certain integral conservation laws that lead to elliptic partial differential equations. In Section 1.3, we detail a three step procedure describing the construction of numerical methods for elliptic equations. In Sections 1.4, 1.5, and 1.6, we compare and contrast three specific numerical methods that arise from our three step procedure: the finite volume method (FV) in Section 1.4, the finite element method (FE) in Section 1.5, and most importantly, the finite volume element method (FVE) in Section 1.6. Finally, a brief outline of the historical development of computational fluid dynamics and roles played by the three methods is given in Section 1.7.

In Chapter 2, we present some theoretical preliminaries. In Sections 2.1 and 2.2, we define and discuss the computational meshes (i.e., the FE triangulation \mathcal{T}^h , the FVE primal mesh $\mathcal{T}^{h/p}$, and the FVE dual mesh $\mathcal{V}^{h/p}$)

that are fundamental to the implementation and analysis of FVE. In Sections 2.3 and 2.4, we define h , p , and h - p version FE and FVE and discuss their relative merits. In Section 2.5, we introduce standard definitions and results for Sobolev spaces that are fundamental to FE and FVE analysis. In Section 2.6, we briefly outline an FE analysis for the steady diffusion equation to set a reference point for an FVE analysis in the next two chapters.

In Chapter 3, we present an h , p , and h - p version FVE analysis for diffusion equations. In Section 3.1, we provide some preliminary definitions and notation that facilitates the transition from FE to FVE analysis. In Section 3.2, we motivate our FVE analysis and its relation to previous FE work. In Section 3.3, we define some discrete norm counterparts to the continuous L^2 and H^1 Sobolev norms of Section 2.5 and demonstrate a discrete Poincaré inequality. As in the FE analysis of Section 2.6, these tools are crucial to FVE analysis. In Section 3.4, we determine an upper bound or continuity result in terms of an FVE diffusion functional and our discrete H^1 semi-norm for the FVE bilinear form that is used to discretize the model diffusion equation discussed in Section 2.6. In Section 3.5, we state and outline applications of the linear Bramble-Hilbert lemma which is at the heart of our FVE analysis. In Section 3.6, we use the Bramble-Hilbert lemma to derive error estimates for the FVE approximation of elements from fractional-order Sobolev spaces by elements from the piecewise polynomial trial spaces employed by the h , p , and h - p version FVE methods: in terms of our FVE diffusion functional, $H^1(\Omega)$ -equivalent optimal-order approximation results are derived for p th-order trial spaces in Section 3.6.1; robust and practical superconvergence results for

linear trial spaces on a “symmetric” volumization are derived in two and three dimensions in Section 3.6.2. In Section 3.7, we derive an “asymptotic” discrete ellipticity result for the FVE diffusion bilinear form in terms of our discrete H^1 semi-norm—aided by the approximation theory from the previous section. Finally in Section 3.8, we construct an FVE analysis for diffusion equations with the components developed in the preceding sections: optimal-order and superconvergence results are obtained for h , p , and h - p version FVE in our discrete H^1 semi-norm. The key results of this chapter are summarized in Section 3.9.

In Chapter 4, we note modifications for an h , p , and h - p version FVE analysis for general elliptic equations and elliptic projections. After presenting some preliminary theoretical infrastructure in Section 4.1, we develop upper bounds in terms of FVE advection and reaction functionals and our discrete L^2 and H^1 norms for the FVE bilinear forms corresponding to the treatment of advection and reaction terms in a general elliptic equation in Section 4.2. In Section 4.3, we develop L^2 -equivalent optimal-order approximation theory results for our FVE advection and reaction functionals. Our goal here is to find estimates for advection and reaction that do not interfere or add any additional “volume symmetry” constraints to the diffusion optimal-order and superconvergence results of the previous chapter. In Section 4.4, we derive “asymptotic” discrete ellipticity results for the FVE reaction and advection bilinear forms. Finally in Section 4.5, we construct an FVE analysis for reaction–diffusion and advection–reaction–diffusion equations with the components developed in the preceding sections: optimal-order and superconvergence results are obtained

for h , p , and h - p version FVE in our discrete H^1 semi-norm. In Section 4.6, we extend the theory developed so far to analyze the elliptic map which is the key theoretical tool in our analysis of parabolic equations. The key results of this chapter are summarized in Section 4.7.

In Chapter 5, we outline an FVE analysis for parabolic equations. First, we define continuous-time and discrete-time numerical methods for general parabolic equations in Section 5.1: the continuous-time method is an FVE variant of the “method of lines”; the discrete-time methods are FVE variants of Backward Euler and Crank-Nicholson “time marching” methods. In Sections 5.2 and 5.3, the components of FVE parabolic analysis are developed and assembled to yield optimal-order and superconvergence results for h , p , and h - p version continuous-time and discrete-time FVE. The key results of this chapter are summarized in Section 5.4 and future work based on the thesis is outlined in Section 5.5.

The main results of the thesis are summarized in Theorems 3.3, 4.3, 4.4, 5.4, and 5.5.

1.2 Background

The finite volume element method for elliptic equations is based on the fact that these equations arise from integral conservation laws [39]. For definiteness and simplicity, we work in two spatial dimensions (modifications to our analysis for one or three dimensions will be stated as needed) and consider the following model elliptic equation for an unknown distribution u :

$$\mathcal{L}u \equiv \nabla \cdot (\mathbf{a}u - D\nabla u) + ru = f \quad \text{in } \Omega, \quad (1.1)$$

where $\mathbf{a} = (a_1, a_2)$, D , and r are advection, diffusion and reaction coefficients, f is a source term, and $\Omega \subset \mathfrak{R}^2$ is the spatial domain. In Chapter 5, we will consider a model parabolic equation based on (1.1) augmented by the mass storage term $m \partial_t u$. For simplicity, we assume here that the coefficients are smooth functions of space alone to avoid difficulties caused by discontinuities and nonlinearities in our analysis, though these cases can be handled (see Section 4.7).

For (1.1), we consider Dirichlet or flux boundary conditions:

$$u = g \quad \text{or} \quad (\mathbf{a}u - D\nabla u) \cdot \mathbf{n} = g \quad \text{on } \partial\Omega, \quad (1.2)$$

where $\partial\Omega$ is the boundary of Ω , g is boundary data on $\partial\Omega$, and \mathbf{n} is the outward unit normal on $\partial\Omega$. To guarantee an unique solution in the case of flux boundary conditions, the source term and boundary data must also satisfy a *compatibility condition*:

$$\int_{\Omega} f \, d\mathbf{x} + \int_{\partial\Omega} g \, dS = 0. \quad (1.3)$$

When coupled with boundary conditions (1.2), the elliptic equation (1.1) is capable of accurately modeling a wide range of physical phenomena in disparate

application areas. Therefore, numerical methods for the elliptic problem defined by equation (1.1) subject to (1.2), like FVE, and their corresponding analysis are vitally important.

Although we stated at the outset that we are investigating numerical solutions of the elliptic partial differential equation (1.1), we are more precisely studying local integral conservation laws on subdomains $V \subseteq \overline{\Omega}$ that lead to equations such as (1.1):

$$\int_{\partial V} (\mathbf{a}u - D\nabla u) \cdot \mathbf{n} \, dS + \int_V r u \, d\mathbf{x} = \int_V f \, d\mathbf{x}, \quad \forall V \subseteq \overline{\Omega}, \quad (1.4)$$

where ∂V is the boundary of V , \mathbf{n} is the outward unit normal on the boundary ∂V , and dS is a spatial boundary measure. The first integral on the left-hand side of equation (1.4) is the net advective-diffusive outflow or outflux of u across ∂V . The second integral on the left-hand side of equation (1.4) is the net internal reaction within V which can be thought of as a type of nonlinear source term. And the single integral on the right-hand side of equation (1.4) is the net contribution due to external sources imposed within V . The local (when $V \subset \overline{\Omega}$) and global (when $V = \overline{\Omega}$) conservation law (1.4) states that the advective-diffusive flux across volume boundary ∂V is counter-balanced by reactions and sources within the volume V . In Chapter 5, we will introduce time dependence and augment (1.4) with the mass storage integral $\int_V m \partial_t u \, d\mathbf{x}$.

Equation (1.4) is called the primitive form of the elliptic equation (1.1) because it contains the essence of the physical model within the least restrictive mathematical model. Finally, it is equation (1.4)—and *not* equation (1.1)—that we hope to mimic discretely in our numerical methods.

1.3 Numerical Methods

FVE is a hybrid method composed from standard finite volume (FV) and finite element (FE) parent methods and is best understood in terms of these methods. Therefore, we will detail a general framework for developing numerical methods that includes FV, FE, and FVE so that the similarities and differences between the methods may be clearly delineated and understood.

When we develop numerical methods, we take the middle ground between the strong form equation (1.1) and the conservation law equation (1.4) and work with the associated variational or weak form equation:

$$\int_{\Omega} w \mathcal{L}u \, d\mathbf{x} = \int_{\Omega} f w \, d\mathbf{x}, \quad \forall w \in \mathcal{W}, \quad (1.5)$$

where \mathcal{W} is a test space of functions defined on the spatial domain Ω . Setting the stage for later developments, we let (\cdot, \cdot) denote the standard $L^2(\Omega)$ inner product and rewrite equation (1.5) in short hand notation:

$$(\mathcal{L}u, w) = (f, w), \quad \forall w \in \mathcal{W}. \quad (1.6)$$

Now the procedure by which we produce numerical methods is easily enumerated in a three step discretization process:

- (1) Discretize the spatial domain Ω in order to
 - (a) define a finite number of sub-domains in Ω on which to pose equation (1.5) and to
 - (b) form a computational mesh whose finite number of degree-of-freedom (DOF) points will define the numerical approximation to the distribution u of equation (1.5);
- (2) Discretize the test space \mathcal{W} of equation (1.5):

- (a) replace \mathcal{W} by a finite dimensional subspace of polynomials test functions defined on the spatial sub-domains of Step 1 so that
 - (b) equation (1.5) is localized on each sub-domain of Step 1;
- (3) Discretize the unknown distribution u and data f of equation (1.5) by
- (a) finite difference or finite element approximations with respect to the DOF points of the computational mesh of Step 1 in order to
 - (b) pose a discrete version of equation (1.5) on each sub-domain of Step 1.

In Step 1, we discretize the spatial domain Ω in two different but interconnected ways: first, we discretize Ω by

- standard finite element triangulations of Ω : \mathcal{T}^h with triangular or rectangular elements denoted by T ; $\mathcal{T}^{h/p}$, the p -fold refinement of \mathcal{T}^h that connects DOF points for polynomials of order p defined each element T in \mathcal{T}^h . Hence, \mathcal{T}^h and $\mathcal{T}^{h/p}$ are basic or primal computational meshes. The details of the triangulation constructions are given in Section 2.1;
- second, we discretize Ω by
- standard finite volume volumizations of Ω , denoted by \mathcal{V}^h or $\mathcal{V}^{h/p}$, with convex or star-shaped volumes, denoted by V , to form dual computational meshes corresponding to the triangulations \mathcal{T}^h and $\mathcal{T}^{h/p}$, respectively. The details of the volumization constructions are given in Section 2.2.

The finite number of degrees-of-freedom located at the vertices the $\mathcal{T}^{h/p}$ triangulation are used represent the numerical approximation to the unknown distribution u of equation (1.5); the set of DOF points on the computational

mesh is denoted by $\{\mathbf{x}_i\}_{i \in I}$, where I is a DOF point indexing set. Each DOF point \mathbf{x}_i of the computational mesh is surrounded by a finite volume, denoted by V_i , so that there is a one-to-one correspondence between DOF points and finite volumes—the specifics of the finite volume construction are given in Section 2.2. Therefore, the set of finite volumes in the volumizations \mathcal{V}^h or $\mathcal{V}^{h/p}$ can be denoted and enumerated by $\{V_i\}_{i \in I}$. According to standard finite volume methodology, the value of the unknown u at DOF point \mathbf{x}_i represents the volume average of u on volume V_i .

In Step 2, we discretize the test space by replacing \mathcal{W} by one of the following choices:

- $\mathcal{S}^p(\mathcal{T}^h)$, the space of piecewise polynomials of order $p \geq 1$ defined on each triangle or rectangle T of the triangulation \mathcal{T}^h ;
- $\mathcal{S}^0(\mathcal{V}^h)$ or $\mathcal{S}^0(\mathcal{V}^{h/p})$, the space of piecewise constants defined on the volumizations \mathcal{V}^h or $\mathcal{V}^{h/p}$, respectively.

In Step 3, we discretize the unknown u , derivatives of u , and the source term f by replacing them with

- finite element approximations u^h , ∇u^h , and f^h , where u^h and f^h are representations of u and f in terms of a finite dimensional trial space chosen to be $\mathcal{S}^p(\mathcal{T}^h)$.
- finite difference approximations to u , ∇u , and f formed by p th-order interpolation and/or extrapolation between vertices of the \mathcal{T}^h triangulation.

Now we can easily describe FV, FE, and FVE by tracing the development of each numerical method through each of the three steps discussed above.

1.4 The Finite Volume Method (FV)

In the finite volume method (FV), we create the computational mesh \mathcal{T}^h and its dual mesh \mathcal{V}^h in Step 1 of Section 1.3. In Step 2 of Section 1.3, we choose $S^0(\mathcal{V}^h)$ as the test space; the set of basis elements for $S^0(\mathcal{V}^h)$ is the set of characteristic functions $\{\chi_i\}_{i \in I}$ (recall that I is a DOF point indexing set) with χ_i defined by

$$\chi_i(\mathbf{x}) \equiv \begin{cases} 1, & \text{if } \mathbf{x} \in V_i, \\ 0, & \text{otherwise.} \end{cases} \quad (1.7)$$

Referring back to equation (1.5), we choose the χ_i basis functions as test functions and apply Green's Theorem (i.e., perform integration by parts) to the advection and diffusion term. As a result, we obtain the FV weak form:

$$\int_{\partial V_i} (\mathbf{a} u - D \nabla u) \cdot \mathbf{n} dS + \int_{V_i} r u d\mathbf{x} = \int_{V_i} f d\mathbf{x}, \quad \forall V_i \in \mathcal{V}^h, \quad (1.8)$$

where ∂V_i is the boundary of V_i , \mathbf{n} is the outward unit normal on the boundary ∂V_i , and dS is a spatial boundary measure. Referring to the short-hand notation of equation (1.6), we perform integration by parts on $(\mathcal{L}u, \chi_i)$ to obtain the left-hand side integrals of equation (1.8). To set the stage for later developments, we rewrite equation (1.8) in the short-hand notation:

$$B(u, \chi_i) = (f, \chi_i), \quad \forall \chi_i \in S^0(\mathcal{V}^h), \quad (1.9)$$

where the the bilinear form $B(\cdot, \cdot)$ on the left-hand side of equation (1.9) consists of all of the integrals on the left-hand side of (1.8).

In Step 3 of Section 1.3, FV uses a somewhat *ad hoc* finite difference methodology that usually corresponds to a reduced or “lumped” quadrature strategy for the integrals of equation (1.8). That is, expressions for u , ∇u , and

f are obtained by interpolation and/or extrapolation between vertex values of u and f in the \mathcal{T}^h triangulation (see Mitchell and Griffiths [36] or Roache [43] for details).

At this point we will defer a discussion about the specifics of obtaining a finite volume numerical solution and the implementation of boundary conditions until we address those issues in conjunction with the finite volume element method.

The principal advantage of the finite volume approach is that its weak form (1.8) closely mimics the integral conservation law (1.4). The method will be locally conservative on each V_i , as well as globally conservative on Ω which is a crucial measure of the physical “correctness” of the numerical solution. Of course, the conservative property does not imply the mathematical correctness of the numerical solution. Usually, we find that the crude finite difference approximations to the integrands of (1.8) limit the effectiveness of FV to relatively simple mesh structures (e.g., Cartesian-product meshes) and simple treatments of boundary conditions.

1.5 The Finite Element Method (FE)

To alleviate some of the difficulties due to FV, one usually turns to the finite element method (FE). For FE, we create the computational mesh \mathcal{T}^h with triangular or rectangular elements and its p -fold refinement $\mathcal{T}^{h/p}$ in Step 1 of Section 1.3. In Step 2 of Section 1.3, we choose $\mathcal{S}^p(\mathcal{T}^h)$ as the test space and denote this choice as \mathcal{W}^h . The set of nodal basis elements for this space \mathcal{W}^h is the set $\{\phi_i\}_{i \in I}$:

$$\phi_i(\mathbf{x}_j) = \delta_{ij}, \quad (1.10)$$

where δ_{ij} is the *Krönecker delta*, and $\phi_i(\mathbf{x})$ is found by p th-order interpolation between the DOF points at the vertices of $\mathcal{T}^{h/p}$. The support of ϕ_i is denoted by \mathcal{T}_i^h and it consists of all elements of the triangulation \mathcal{T}^h that have \mathbf{x}_i as a vertex.

Referring back to equation (1.5), we choose the ϕ_i basis functions as test functions and perform integration by parts on the advection and diffusion terms. As a result, we obtain the FE weak form given by (1.11) on sub-domains away from the domain boundary $\partial\Omega$:

$$\int_{\mathcal{T}_i^h} D\nabla u \cdot \nabla \phi_i - u \mathbf{a} \cdot \nabla \phi_i + r u \phi_i d\mathbf{x} = \int_{\mathcal{T}_i^h} f \phi_i d\mathbf{x}, \quad \forall \phi_i \in \mathcal{W}^h, \quad (1.11)$$

where \mathcal{T}_i^h is the support of ϕ_i . To describe modifications on the domain boundary $\partial\Omega$, we define $\partial\Omega_i$ as the intersection of the domain boundary with $\overline{\mathcal{T}_i^h}$ —the closure of the i th test function support: $\partial\Omega_i \equiv \partial\Omega \cap \overline{\mathcal{T}_i^h}$. When $\partial\Omega_i$ has positive measure, the FE weak form (1.11) includes the boundary flux integral

$$\int_{\partial\Omega_i} (\mathbf{a} u - D\nabla u) \cdot \mathbf{n} \phi_i dS. \quad (1.12)$$

Regarding boundary modifications, let us pause for a few comments regarding implementation of the boundary conditions specified in (1.2) with respect to the boundary integral of (1.12)—additional details can be found in the texts of Ciarlet [11] or Johnson [28]. If a Dirichlet boundary condition is imposed, the boundary flux integrand of (1.12) is an unknown and can be solved for separately if desired. The information from the Dirichlet boundary condition is imposed directly upon the trial space discretization of u on $\partial\Omega$ and information from boundary nodes on $\partial\Omega$ are passed to the right-hand side of our discrete equations (cf. equation (1.16) below) in an obvious manner that is discussed in the texts cited above. However, if a flux boundary condition is imposed, the boundary flux is known data and boundary flux integral of (1.12) is moved to the right-hand side of equation (1.11). In the discussion below, we'll assume that a homogeneous flux boundary condition has been imposed in order to simplify the presentation.

Referring to the short-hand notation of equation (1.6), we perform integration by parts on $(\mathcal{L}u, \phi_i)$ to obtain the left-hand side integrals of equation (1.11) and the boundary integral of (1.12). To set the stage for later developments, we rewrite equation (1.11) in short-hand notation:

$$A(u, \phi_i) = (f, \phi_i), \quad \forall \phi_i \in \mathcal{W}^h, \quad (1.13)$$

where the bilinear form $A(\cdot, \cdot)$ on the left-hand side of equation (1.13) consists of all of the integrals on the left-hand side of (1.11).

In Step 3 of Section 1.3, the Galerkin FE uses \mathcal{W}^h as a trial space, as well as a test space, to represent the unknown u and source f of equation (1.11), but other higher order polynomial spaces defined on the mesh can be used as

a trial space and the result is termed Petrov-Galerkin FE. In the remainder of this dissertation, we'll refer to the Galerkin FE as “the” finite element method since it is the FE of choice in the bulk of applications. To describe our finite element method, we define $u^h \in \mathcal{W}^h$ as the trial space representation of the numerical solution:

$$u^h \equiv \sum_{j \in I} u_j^h \phi_j(\mathbf{x}), \quad (1.14)$$

where $u_j^h = u^h(\mathbf{x}_j)$ so that u^h has a nodal representation. Similarly, we define $f^h \in \mathcal{W}^h$ as our trial space representation of the source term:

$$f^h \equiv \sum_{j \in I} f_j \phi_j(\mathbf{x}), \quad (1.15)$$

where $f_j = f(\mathbf{x}_j)$ so that $f^h = \Pi_h^p f$ is the *p*th-order interpolant of f into $\mathcal{S}^p(\mathcal{T}^h)$.

Referring to equation (1.13), the finite element method solution of equation (1.1) is described as follows: find $u^h \in \mathcal{W}^h$ such that

$$A(u^h, \phi_i) = (f^h, \phi_i), \quad \forall \phi_i \in \mathcal{W}^h. \quad (1.16)$$

Equivalently, the simultaneous solution to the finite dimensional problem given by equation (1.16) is given by the corresponding matrix problem:

$$\mathbf{A} \mathbf{u} = \mathbf{f}_{\text{FE}}, \quad (1.17)$$

where the finite element matrix \mathbf{A} of (1.17) has entries $[\mathbf{A}]_{ij}$ defined by

$$[\mathbf{A}]_{ij} \equiv A(\phi_j, \phi_i), \quad (1.18)$$

and the vector \mathbf{u} of (1.17) contains the nodal values of the finite element solution and has entries $[\mathbf{u}]_j$ given by

$$[\mathbf{u}]_j \equiv u_j^h. \quad (1.19)$$

As a result of the definitions of u^h in (1.14), \mathbf{A} in (1.18), and \mathbf{u} in (1.19), we see that the i th entry of the vector formed by the matrix/vector inner product on the left-hand side of equation (1.17), $[\mathbf{A}\mathbf{u}]_i$, is equivalent to the left-hand side of equation (1.16) for the i th test function:

$$[\mathbf{A}\mathbf{u}]_i = \sum_{j \in I} u_j^h A(\phi_j, \phi_i) = A(u^h, \phi_i). \quad (1.20)$$

Similarly, we define the i th entry of the finite element right-hand side vector of (1.17)— $[\mathbf{f}_{\text{FE}}]_i$ —as the right-hand side of equation (1.16) for the i th test function: recalling the definition of f^h in (1.15), we have

$$[\mathbf{f}_{\text{FE}}]_i \equiv (f^h, \phi_i) = \sum_{j \in I} f_j(\phi_j, \phi_i). \quad (1.21)$$

The principal advantage of the finite element approach is its mathematical generality: FE triangulations can be fitted to very general geometries and boundary conditions; the trial and test spaces defined on those triangulations provide accurate and robust approximations to the unknown u and data f . However, the mathematical generality of FE can be a weakness as well as a strength: the computational overhead required to set up the FE matrix system of equations can be expensive compared to the set up costs for FV on general computational meshes. Furthermore, this overhead is relatively fixed: the set up cost for FE is largely problem independent for a general mesh—relatively simple diffusion problems require nearly as much work as complex advection–reaction–diffusion problems. From the physical standpoint, the boundary integral (1.12) implies that FE will match the form of the governing conservation law globally on Ω but not on any sub-domain of Ω ; hence an important characteristic of the problem—local conservation—has been lost.

1.6 The Finite Volume Element Method (FVE)

To realize the local conservation property of the finite volume method within a finite element framework, we turn to the finite volume element method (FVE)—see Baliga and Patankar [4] and McCormick [39] for fundamentals. In Steps 1 and 2 of Section 1.3, we follow the a “combined” FE and FV methodology: in Step 1, we construct a FE triangulation \mathcal{T}^h with triangular or rectangular elements, its p -fold refinement $\mathcal{T}^{h/p}$, and a dual mesh $\mathcal{V}^{h/p}$ that is tied to $\mathcal{T}^{h/p}$ rather than \mathcal{T}^h as in FV—the importance of this distinction will be discussed below; in Step 2, choose $\mathcal{S}^0(\mathcal{V}^{h/p}) = \mathcal{X}^h$ as a test space. After integration by parts, FVE has the same weak form as FV—cf. equations (1.8) and (1.9)—which preserves a finite dimensional analog of the local conservation property of the integral conservation law (1.4). However in Step 3 of Section 1.3, we break away from the usual FV discretization of the unknown u and the data f and turn to the usual FE discretization: replace u and f with representations u^h and f^h in the trial space $\mathcal{S}^p(\mathcal{T}^h) = \mathcal{W}^h$ as in equations (1.14) and (1.15). We can do this because the volumization $\mathcal{V}^{h/p}$ establishes a one-to-one correspondence between the vertices of $\mathcal{T}^{h/p}$ which are the DOF points for $\mathcal{S}^p(\mathcal{T}^h)$. Hence, the use of the FE trial space is equivalent to systematic and generalized FV-style interpolations or extrapolations between nodes of $\mathcal{T}^{h/p}$.

Regarding boundary conditions, the general remarks for the finite element method also apply to the finite volume element method. As in the FE development of the previous section, we’ll assume a homogeneous flux boundary to simplify our presentation. Further discussion of boundary conditions and their impact on our FVE analysis will be deferred to later chapters.

Referring to equation (1.9), the finite volume element method solution for equation (1.1)—or perhaps more appropriately equation (1.4)—is described as follows: find $u^h \in \mathcal{W}^h$ such that

$$B(u^h, \chi_i) = (f^h, \chi_i), \quad \forall \chi_i \in \mathcal{X}^h, \quad (1.22)$$

where f^h satisfies an L^2 projection relationship with f ,

$$(f^h, \chi_i) = (f, \chi_i), \quad \forall \chi_i \in \mathcal{X}^h, \quad (1.23)$$

to ensure that the numerical source term conserves mass on each volume in $\mathcal{V}^{h/p}$. Equivalently, the simultaneous solution to the finite dimensional problem given by equation (1.22) is given by the corresponding matrix problem:

$$\mathbf{B} \mathbf{u} = \mathbf{f}_{\text{FVE}}, \quad (1.24)$$

where the finite volume element matrix \mathbf{B} of (1.24) has entries $[\mathbf{B}]_{ij}$ defined by

$$[\mathbf{B}]_{ij} \equiv B(\phi_j, \chi_i), \quad (1.25)$$

and as before the vector \mathbf{u} of (1.24) contains the nodal values of the finite volume element solution and has entries $[\mathbf{u}]_j$ of the form given by (1.19). As a result of the definitions of u^h in (1.14), \mathbf{B} in (1.25), and \mathbf{u} in (1.19), we see that the i th entry of the vector formed by the matrix/vector inner product on the left-hand side of equation (1.24) is equivalent to the left-hand side of equation (1.22) for the i th test function:

$$[\mathbf{B}\mathbf{u}]_i = \sum_{j \in I} u_j^h B(\phi_j, \chi_i) = B(u^h, \chi_i). \quad (1.26)$$

Similarly, we define the i th entry of the finite volume element right-hand side vector of (1.24)— \mathbf{f}_{FVE} —as the right-hand side of equation (1.22) for the i th test function: recalling the definition of f^h in (1.15), we have

$$[\mathbf{f}_{\text{FVE}}]_i \equiv (f^h, \chi_i) = \sum_{j \in I} f_j(\phi_j, \chi_i). \quad (1.27)$$

The principal advantage of the finite volume element approach is that its FV weak form (1.8) mimics the integral conservation law (1.4) while incorporating the mathematical generality of FE for very general geometries and boundary conditions. From the FV perspective, FVE represents a systematic and general approach to generating finite difference stencils that is only slightly more complex than the standard FV approach. While from the FE perspective, FVE represents a locally conservative variant of the standard FE by the choice of a simpler, lower order, volume-oriented test space.

From the standpoint of application, FVE favors its FV parent as it agrees with FV in Steps 1 and 2 of its formation. In Step 3, FVE can build upon FV by using a FE trial space rather than interpolate or extrapolate between nodes of \mathcal{T}^h . From the standpoint of analysis, FVE favors its FE parent as it can be viewed as a Petrov-Galerkin variant of the standard finite element method. In classical finite volume analysis (see Mitchell and Griffiths [36], Kreiss *et al.* [31], Manteuffel and White [35], and the references listed therein) FV is viewed as a special type of finite difference method and analyzed in terms of finite difference truncation error in C^∞ norms. However, in modern finite volume analysis (see [5, 17, 18, 21, 23, 32, 33, 38, 46, 50, 54]) FV is viewed as a special type of finite element method and its error is analyzed in terms of finite element approximation theory results: results for the error in approximating

elements of a fractional-order Sobolev space with piecewise polynomials. Since FVE is even closer to FE than FV, our analysis will follow that of FE even more closely.

Once we have established some theoretical infrastructure in the next chapter, we'll come back to the theoretical similarities between FVE and FE. But for now, we need to close the book with regard to the basic description of the methods by briefly discussing the historical development of the computational fluid dynamics and the roles played by FV, FE, and FVE.

1.7 Historical Development

Historically (see Batchelor [6]), the local and global conservation law (1.4) was the starting point for theoretical fluid dynamics in the 18th century; later with the development of continuum mechanics, it was shown (under suitable conditions) that (1.4) implied (1.1). In the current century (see Roache [43], Shashkov and Steinberg [48]), finite difference numerical methods based on (1.4) have enjoyed great success in a wide range of application problems in fluid dynamics.

Starting in the 19th century (see Guenther and Lee [20]) with the development of the calculus of variations and functional analysis, attention in mathematical and physical circles shifted from local conservation laws to integral-based variational principles that (under suitable conditions) lead to equations like (1.1) in fluid dynamics as well as in other branches of continuum mechanics. In the current century (see Johnson [28] and Oden [40]), finite element numerical methods (discussed in Section 2.6) based on integral variational principles have succeeded in cases where finite differences have encountered difficulties and even failures.

Coupled with a sophisticated yet elegant analysis (see [11, 12, 13, 14, 55]) that is one of the crowning jewels of modern applied mathematics, it appeared likely that Galerkin finite element methods would largely supplant finite differences by the end of this century. However, this has not happened in fluid dynamics and is not likely to happen in the future. Somewhat superficially, the issue of computational efficiency has been problematic: the computational

set-up costs for finite elements are relatively expensive when compared to finite differences on general computational grids that discretize domains with complex geometries.

More fundamentally, the problem lies in the choice of the underlying model for computation. Standard finite volume or cell-centered finite difference methods (FV) rely on the local and global conservation law (1.4), which is always physically valid even when extrapolation to (1.1) is not. Standard Galerkin finite element methods (FE) rely on an integral formulation of (1.1) arising from variational principles and are not locally conservative; however, FE is globally conservative. Hence, FE violate the spirit and the letter of the fundamental local conservation laws that govern fluid dynamics.

Furthermore, the difficulties that arise in FV are not due to its conservation law model but are usually due to somewhat *ad hoc* discretizations of the unknown solution, known data, and boundary conditions for problems posed on difficult geometries. Conversely, the successes of FE are largely due to its systematic discretizations based on piecewise polynomial spaces defined on FE triangulations for general geometries and not due to its variational model.

At the same time, variational principles are fundamental (and conservation laws are not) in related disciplines of continuum mechanics such as structural mechanics and electrostatics which are based on a principle of “virtual work.” In those applications, FE is the preferred numerical method (see Babuška and Osborn [2] and the references cited therein).

In FVE, we try to combine the best of FV and FE. First, it is the local and global conservation law (1.4)—and *not* a FE weak form of (1.1)—that

we mimic discretely in FVE. Second, FVE approximates (1.4) using meshes described in the next two sections, by replacing u and f in (1.4) with finite element approximations u^h and f^h which are based on a finite element triangulation \mathcal{T}^h that partitions Ω , and by posing equation (1.4) on a finite subset $\mathcal{V}^{h/p}$ of volumes that partitions Ω .

FVE is a full-fledged finite element method that is based on a local conservation model rather than a variational model; but at the same time, FVE can recover more computationally efficient equations like those of FV with the use of certain reduced or lumped quadratures (see [5]). In general, FVE enjoys some of the computational advantages of FV when compared to FE. In FE, piecewise polynomial trial and test spaces of degree $p \geq 1$ are used to discretize the unknown u and to set-up FE equations, respectively. Therefore, quadrature rules that are exact for polynomials of degree $2p$ are used to evaluate multiple integrals in FE. In FVE, piecewise polynomial trial spaces of degree p and piecewise polynomial test spaces of degree 0 are used to discretize the unknown u and to set-up FVE equations, respectively. Therefore, quadrature rules that are exact for polynomials of degree p are used to evaluate multiple integrals in FVE. If we let d be the number of spatial dimensions in a given problem domain, we see that FVE quadrature will be $O(2^d)$ times less expensive than FE quadrature for a given problem (see Stroud [49] for a discussion of quadrature for multiple integrals). For practical engineering problems in three dimensions, the computational set-up costs for FVE are nearly an order of magnitude less expensive than FE.

1.8 Summary

In this chapter, we have given the necessary introduction for the finite volume element method (FVE) for elliptic equations as a prelude to an analysis of the method. In Section 1.2, we outlined how FVE arose from the approximation of certain integral conservation laws that led to elliptic partial differential equations. In Section 1.3, we detailed a three step procedure for the construction of numerical methods for elliptic equations. In Sections 1.4, 1.5, and 1.6, we compared and contrasted three specific numerical methods that arose from our three step procedure: the finite volume method (FV) in Section 1.4, the finite element method (FE) in Section 1.5, and most importantly, the finite volume element method (FVE) in Section 1.6. Finally, a brief outline of the historical development of computational fluid dynamics and roles played by the three methods was given in Section 1.7.

2. Theoretical Preliminaries

In this chapter, we make the discussion started in the previous chapter more precise by providing some theoretical preliminaries: construction of the FVE triangulations and volumizations, definitions of h , p , and h - p version FVE, a comparison between FE and FVE versions, Sobolev spaces to characterize the exact solution to the model elliptic equation (1.1), and an outline of standard FE analysis.

2.1 Triangulation

The most basic component in the implementation and analysis of the finite volume element method is the discretization of the domain Ω into computational meshes. The finite element triangulation \mathcal{T}^h of Ω and the finite volume element volumization $\mathcal{V}^{h/p}$ of Ω are two different—yet interconnected—meshes or discretizations of Ω . The two discretizations are connected by a third mesh or discretization of Ω : $\mathcal{T}^{h/p}$, the p -fold refinement of \mathcal{T}^h . The nodes or vertices of $\mathcal{T}^{h/p}$ are the locations of the degrees of freedom (DOF) for $u^h \in \mathcal{S}^p(\mathcal{T}^h)$: u^h is a C^0 function on Ω that is a polynomial of degree $p \geq 1$ when restricted to each element T of \mathcal{T}^h . In FVE terminology, the triangulation $\mathcal{T}^{h/p}$ is the *primal* mesh and the volumization $\mathcal{V}^{h/p}$ is the *dual* mesh. Next, we describe these three meshes and their relationships.

2.1.1 FE Triangulation

Let \mathcal{T}^h be a non-overlapping triangulation of $\overline{\Omega}$, the closure of Ω , into a finite number of elements T . That is, \mathcal{T}^h satisfies the following general properties [11]:

- (1) For each $T \in \mathcal{T}^h$, the set T is closed and its *interior*, $\text{int}(T)$, is nonempty and connected;
- (2) For each $T \in \mathcal{T}^h$, the boundary ∂T is Lipschitz-continuous;
- (3) \mathcal{T}^h partitions $\overline{\Omega}$: i.e., $\overline{\Omega} = \bigcup_{T \in \mathcal{T}^h} T$.
- (4) The elements of \mathcal{T}^h *do not overlap*: for each distinct $T_1, T_2 \in \mathcal{T}^h$, we have $\text{int}(T_1) \cap \text{int}(T_2) = \emptyset$.
- (5) Any face of any $T_1 \in \mathcal{T}^h$ is either a subset of the boundary $\partial\Omega$ or a face of another $T_2 \in \mathcal{T}^h$.

To simplify the discussion, we assume the elements of the triangulation are triangles (possibly curved to accommodate non-polygonal domains). Other types of elements could be considered: in particular, rectangular elements for a rectangular domain Ω .

For each T of \mathcal{T}^h , we define the mesh parameters h_T, h , and ρ_T as follows: h_T is the diameter of the circumscribing circle for T ; h is the maximum value of h_T ; and ρ_T is the diameter of the inscribed circle in T .

We assume that the triangulation is shape *regular*: there exists a positive constant σ such that

$$\frac{h_T}{\rho_T} \leq \sigma, \quad \forall T \in \mathcal{T}^h; \tag{2.1}$$

the family (h_T) is bounded and 0 is its unique accumulation point—i.e., h approaches zero. The assumption of regularity is used to simplify error estimates and to avoid degenerate triangulations (i.e., a minimum angle condition is satisfied). See Figure 2.1 for an example of the triangulation \mathcal{T}^h .

2.1.2 FVE Primal Mesh

Once a nodal or base triangulation \mathcal{T}^h is defined, we can define elements of $\mathcal{S}^p(\mathcal{T}^h)$. The degrees of freedom (DOF) for $u^h \in \mathcal{S}^p(\mathcal{T}^h)$ are located in a regular fashion in \bar{T} : on the vertices, along edges, and in the interior of T . Just as the FE triangulation \mathcal{T}^h can be determined by connecting the vertices of T to their nearest neighbors, an alternative triangulation $\mathcal{T}^{h/p}$ can be determined by connecting the DOF in T to their nearest neighbors within T : let T_p denote the triangular elements of $\mathcal{T}^{h/p}$. See Figure 2.2 for an example of the triangulation $\mathcal{T}^{h/p}$ restricted to a single element in \mathcal{T}^h .

To be specific, $\mathcal{T}^{h/p}$ is generated from \mathcal{T}^h by the “newest vertex bisection” algorithm developed by Sewell [47] and applied by Mitchell [37] to define mesh hierarchies and the corresponding hierarchical basis functions used in the Multigrid iterative solution of p -version FE problems. This refinement strategy is very similar to the FAC (Fast Adaptive Composite grid) algorithms that were successfully applied to h -version FVE in the monograph of McCormick [39]. Although we will not discuss the iterative solution of p -version FVE problem, we use this mesh construction to define a volume mesh in Section 2.2.1, and to define p and h - p version FVE in Section 2.3. At the same time, we wish to set the stage for the efficient numerical implementation p and h - p version FVE at a later date.

Figure 2.1: An example of the triangulation \mathcal{T}^h .

Figure 2.2: $\mathcal{T}^{h/p}$ restricted to a single element of \mathcal{T}^h .

Note that each element in $\mathcal{T}^{h/p}$ is contained in a single element of \mathcal{T}^h . The primal mesh $\mathcal{T}^{h/p}$ can be seen as the p -fold refinement of \mathcal{T}^h : if $\text{diam}(T) = h$, then $\text{diam}(T_p) = h/p$. Therefore, we write

$$\mathcal{T}^{h/p} \subseteq \mathcal{T}^h \tag{2.2}$$

as a shorthand notation for this mesh hierarchy or nested property of the triangulations.

2.2 Volumization

The FVE dual mesh $\mathcal{V}^{h/p}$ or volumization of Ω partitions $\overline{\Omega}$ into a finite number of non-overlapping elements V_i —the index i refers to a one-to-one correspondence between volumes and DOF for $u^h \in \mathcal{S}^p(\mathcal{T}^h)$. To define $\mathcal{V}^{h/p}$ from the primal mesh $\mathcal{T}^{h/p}$, we make frequent reference to $\mathcal{T}_i^{h/p}$: the union of all $T_p \in \mathcal{T}^{h/p}$ that have the location of the i th DOF, \mathbf{x}_i , as a vertex. In general, $\mathcal{V}^{h/p}$ satisfies the following properties:

- (1) For each $V \in \mathcal{V}^{h/p}$, the set V is closed and its *interior*, $\text{int}(V)$, is nonempty and connected;
- (2) For each $V \in \mathcal{V}^{h/p}$, the boundary ∂V is Lipschitz-continuous;
- (3) $\mathcal{V}^{h/p}$ partitions $\overline{\Omega}$: i.e., $\overline{\Omega} = \bigcup_{V \in \mathcal{V}^{h/p}} V$.
- (4) The elements of $\mathcal{V}^{h/p}$ *do not overlap*: For each distinct $V_i, V_j \in \mathcal{V}^{h/p}$, one has $\text{int}(V_i) \cap \text{int}(V_j) = \emptyset$.
- (5) Any face of any $V_i \in \mathcal{V}^{h/p}$ is either a subset of the boundary $\partial\Omega$ or a face of another $V_j \in \mathcal{V}^{h/p}$;
- (6) There is a one-to-one correspondence between primal mesh nodes and dual mesh volumes: $\mathbf{x}_i \in V_i$ and $V_i \subset \mathcal{T}_i^{h/p}$, $\forall i \in I$.

Properties 1 through 5 are satisfied by any triangulation. However, Property 6 distinguishes the volumization $\mathcal{V}^{h/p}$ from triangulation $\mathcal{T}^{h/p}$ and hints at the concrete relationship between the FVE primal and dual meshes that will be detailed below.

2.2.1 Volume Construction

To make our discussion more concrete, we follow the general volume construction algorithm of Bank and Rose [5]:

- (1) Select a point $P \in \overline{T_p}$, $\forall T_p \in \mathcal{T}_i^{h/p}$;
- (2) Connect P by straight line segments to edge midpoints of T_p for the two edges of T_p adjacent to the vertex \mathbf{x}_i , $\forall T_p \in \mathcal{T}_i^{h/p}$;
- (3) For each $T_p \in \mathcal{T}_i^{h/p}$, define a *sub-volume*, $v_i(T_p)$, as the region bounded by the line segments formed in Step 2 and line segments connecting the edge midpoints of T_p with the vertex \mathbf{x}_i . Note that $v_i(T_p) = V_i \cap T_p$; see Figure 2.3.

Finally, define the volume V_i as the region enclosed by all sub-volumes formed in Step 3:

$$V_i \equiv \bigcup_{T_p \in \mathcal{T}_i^{h/p}} v_i(T_p). \quad (2.3)$$

Let $v^{h/p}$ denote the set of all sub-volumes; then, $v^{h/p}$ is a refinement of or nested in $\mathcal{V}^{h/p}$:

$$v^{h/p} \subseteq \mathcal{V}^{h/p}. \quad (2.4)$$

More importantly, each sub-volume $v_i(T_p)$ of $v^{h/p}$ is contained in a single element T_p of $\mathcal{T}^{h/p}$; therefore, $v^{h/p}$ is nested in $\mathcal{T}^{h/p}$ and in \mathcal{T}^h :

$$v^{h/p} \subseteq \mathcal{T}^{h/p} \subseteq \mathcal{T}^h. \quad (2.5)$$

In our FVE analysis in Chapter 4, $v^{h/p}$ will be the focus of a local error analysis for interpolation defined on elements of $\mathcal{T}^{h/p}$ and then restricted to elements of $v^{h/p}$ on an element-by-element basis. Due to the nesting of the sub-volumes in the triangulations, we obtain error estimates on \mathcal{T}^h by summing over results on $v^{h/p}$.

Figure 2.3: Construction of the sub-volume $v_i(T_p)$.

The choice of P in Step 1 is crucial in volume construction. In practice (see Cai [10]), we typically use one of the following volume types (characterized by P):

- (1) Circumcenter volume: P is the center of the circle circumscribed about T_p ; equivalently, P is the point of intersection for the perpendicular bisectors of the edges of T_p . To ensure that $P \in \overline{T_p}$, we require that no interior angle of T_p exceeds $\pi/2$ radians—no obtuse triangles are permitted.
- (2) Centroid volume: P is the center of mass or centroid of T_p ; equivalently, P is the point of intersection of the medians of T_p . A *median* of T_p is a line segment connecting a vertex of T_p with the midpoint of its opposite side.
- (3) Incenter volume: P is the center of the circle inscribed in T_p ; equivalently, P is the point of intersection of the angle bisectors of T_p .
- (4) Orthocenter volume: P is the point of intersection of the altitudes of T_p . An *altitude* of T_p is the line segment from a vertex of T_p that is perpendicular to the line containing the opposite side. To ensure that $P \in \overline{T_p}$, we require that no interior angle of T_p exceeds $\pi/2$ radians—no obtuse triangles are permitted.

In applications, advantages of circumcenter volumes are that they are always convex and geometrically simple (see Figure 2.4 for an example), while other types of volumes are usually non-convex and geometrically complex (see Figure 2.5 for an example and Bank and Rose [5] for details).

Figure 2.4: A circumcenter finite volume V_i .

Figure 2.5: A general finite volume V_i .

2.2.2 Volume Symmetry

In the analysis to follow, superconvergence results for linear finite elements can be demonstrated for “symmetric” circumcenter volumes. To define this symmetry precisely, we need additional notation: X_{ij} is the edge or line segment connecting nodes \mathbf{x}_i and \mathbf{x}_j of the triangulation $\mathcal{T}^{h/p}$; γ_{ij} is the interface or volume boundary between volumes V_i and V_j of the volumization $\mathcal{V}^{h/p}$ —i.e., $\gamma_{ij} \equiv \overline{V_i} \cap \overline{V_j}$. Following Cai *et al.* [9] and Cai [10], the volumization $\mathcal{V}^{h/p}$ is *symmetric* to the triangulation $\mathcal{T}^{h/p}$ if the following two symmetries hold for all volumes in $\mathcal{V}^{h/p}$:

- (X -symmetry) γ_{ij} is a perpendicular bisector of X_{ij} ;
- (γ -symmetry) X_{ij} is a perpendicular bisector of γ_{ij} .

We remark that for rectangular elements, only the first condition of X -symmetry is required for $\mathcal{V}^{h/p}$ to be symmetric to $\mathcal{T}^{h/p}$ (cf. Süli [50]).

2.2.3 Volume Regularity

Of more general significance is the notion of volume “regularity,” which is assumed in all of our results (cf. Cai [10]). Form an auxiliary triangulation $\widetilde{\mathcal{T}}^{h/p}$ of triangular elements by connecting the endpoints of γ_{ij} with the endpoints of X_{ij} for every volume interface in $\mathcal{V}^{h/p}$. That is, we define a “macro-element” E_{ij} as the region formed by joining the endpoints of γ_{ij} and X_{ij} (See Figure 2.6 for an illustration). Then E_{ij} is the union of two triangular elements (with X_{ij} as a common edge): i.e, $E_{ij} = T_{ij}^1 \cup T_{ij}^2$, where each element $T_{ij}^k, k \in \{1, 2\}$, is contained in a single element of $\mathcal{T}^{h/p}$. Therefore, $\widetilde{\mathcal{T}}^{h/p}$ is a refinement of $\mathcal{T}^{h/p}$ and $\text{diam}(T_{ij}^k) = O(h/p)$ and we have the following nesting

of triangulations:

$$\widetilde{\mathcal{T}}^{h/p} \subseteq \mathcal{T}^{h/p} \subseteq \mathcal{T}^h. \quad (2.6)$$

In our FVE analysis in Chapters 3 and 4, $\widetilde{\mathcal{T}}^{h/p}$ will be the focus of a local error analysis for interpolation defined on elements of $\mathcal{T}^{h/p}$ and then restricted to elements of $\widetilde{\mathcal{T}}^{h/p}$ on an element-by-element basis. Due to the nesting of the triangulations, we obtain error estimates on \mathcal{T}^h by summing over results on $\widetilde{\mathcal{T}}^{h/p}$.

For circumcenter volumes (the focus of superconvergent error estimates in Chapter 3), if this auxiliary triangulation $\widetilde{\mathcal{T}}^{h/p}$ is shape regular according to the definition in Section 2.1.1 (i.e., inequality (2.1) is satisfied) and no interior angle of $\widetilde{T}_p \in \widetilde{\mathcal{T}}^{h/p}$ exceeds $\pi/2$, then the volumization $\mathcal{V}^{h/p}$ is *regular*. For the other volume types of Section 2.2.1, shape regularity of $\widetilde{\mathcal{T}}^{h/p}$ implies regularity of $\mathcal{V}^{h/p}$. See Figure 2.6.

Figure 2.6: Construction of a pair of elements in $\widetilde{\mathcal{T}}^{h/p}$.

2.3 Versions h , p , and h - p

Motivated by the success of the h , p , and h - p version FE of Babuška in structural mechanics problems, we define and analyze h , p , and h - p FVE for fluid dynamics problems under uniform refinements of h and p in direct analogy to the FE development outlined in [2, 3, 27, 51].

In the h -version FVE, we fix the order p of the FE polynomial trial space and uniformly refine the computational mesh \mathcal{T}^h . The monograph of McCormick details mesh refinement and the efficient numerical solution of h -version FVE discretizations. For $p = 1$, the paper of Cai and McCormick [8] analyzes local refinement of the h -version FVE for a steady diffusion equation; similarly, Ewing *et al.* [17, 18] and Lazarov *et al.* [32] analyze local refinement for h -version FV for general elliptic and parabolic problems.

In the p -version FVE, we fix the computational mesh \mathcal{T}^h and uniformly increase the order p of our FE trial space, which in turn refines the FVE meshes $\mathcal{T}^{h/p}$ and $\mathcal{V}^{h/p}$. In the h - p version FVE, we refine uniformly in h and p . As in p and h - p version FE (see Szabo and Babuška [51] for background), hierarchical basis functions could be used which induce a mesh hierarchy for $\mathcal{T}^{h/p}$ and $\mathcal{V}^{h/p}$ in the p and h - p version FVE as discussed in Mitchell [37]. However, our focus here is on an *a priori* error analysis, so it is sufficient that we have a set of basis functions (not necessarily hierarchical) on each element of \mathcal{T}^h that is defined in terms of the vertex DOF on $\mathcal{T}^{h/p}$. So for simplicity, we can assume that a standard (non-hierarchical) nodal basis is being used. However, in an *a posteriori* error analysis and in real-world implementations, hierarchical basis functions are essential to an efficient iterative solution (see Hu and Katz [25],

Mandel [34], Mitchell [37], and Pavarino [42] for background). To this end, we note that our definitions of $\mathcal{T}^{h/p}$ and $\mathcal{V}^{h/p}$ in the previous two sections were stated with efficient implementation in mind and fully support basis hierarchy according to the “newest vertex bisection” algorithm discussed in [37].

Although our error analysis in the remainder of the text is independent of any particular solution algorithm for FVE discretizations, we are motivated and strongly influenced by the success of Multigrid and Multilevel iterative algorithms (MG) in numerical partial differential equations (see McCormick [39] and the references cited therein for background). Out of the standard h -version FV/FD, FE, and FVE discretizations, McCormick [39] found that FVE was the most amenable to efficient MG solutions. Since the success of MG and Domain Decomposition solution algorithms [25, 34, 37, 42] for p and h - p version FE discretizations is either directly based on or closely related to the ideas presented in [39], we can reasonably expect similar success for p and h - p version FVE. In fact, it was the success of these methods for FE that prompted the analysis of p and h - p version FVE in this thesis. If the performance of FVE vs. FE for p and h - p methods parallels the MG performance of the h -version methods, FVE may be superior. Some anecdotal support for this position was stated in Section 1.2: an operation count for the costs of equation set-up in FVE vs. FE (i.e., numerical quadrature of multiple integrals) revealed that FVE was nearly an order of magnitude less expensive for three dimensional problems. Further support for this position is given in the next section.

2.4 FVE vs. FE

In Sections 1.1 and 1.2, we outlined the origin of fluid dynamics and its basis on local conservation laws. Also, we motivated the need for FVE which preserves local conservation on the discrete level. In addition, we saw that FVE systematically discretizes the unknown solution, known source term data and boundary conditions on domains with complex geometries via the FE triangulation \mathcal{T}^h , the FVE triangulation $\mathcal{T}^{h/p}$, and the FVE volumizations $\mathcal{V}^{h/p}$. From this perspective, we can view FVE as a locally conservative Petrov-Galerkin FE. This prompted us to define h , p , and h - p version FVE in direct analogy to h , p , and h - p FE in the previous section. In addition to local conservation and the variety of versions, FVE has numerous other practical advantages over FE that are outlined below.

In Section 1.2, we saw that the computational set-up costs for FVE were less expensive than the costs for Galerkin FE. If d is the spatial dimension of the problem domain, the quadrature rules used to evaluate FVE integrals and set-up a system of linear equations were seen to be $O(2^d)$ less expensive than corresponding FE quadrature rules. That is, the quadrature operation count for reaction and source term integrals is $O(p)^d$ for FVE and $O(2p)^d$ for FE, since the order of polynomial trial and test space is p and 0 for FVE and p and p for FE.

Furthermore, for reaction and source terms, the FVE integrals from (1.4) are defined on volumes from $\mathcal{V}^{h/p}$ with diameters of $O(h/p)$ (see Section 2.2), whereas FE integrals are defined on elements from \mathcal{T}^h with diameters of $O(h)$ (see Section 2.1); therefore, the FVE quadratures are performed on

regions that are $O(p)^d$ smaller than their FE counterparts. This feature makes the FVE equations more local in character and the FVE mass matrix arising from the reaction term more strongly diagonally dominant than its FE counterpart; from the standpoint of numerical linear algebra and implementation, this means that it should be easier to precondition and iteratively solve FVE systems of linear equations than their FE counterparts.

In addition, FVE diffusive and advective fluxes are defined as surface integrals in (1.4) which are simpler (i.e., of lower dimension) than the corresponding volume integrals for FE: i.e., the quadrature operation count is $O(p)^{d-1}$ for FVE and $O(2p)^d$ for FE—a increased savings of $O(p2^d)$. Coupled with the documented [39] success of Multi-level solution algorithms for h -version FVE vs. FV discretizations and the documented [25, 35, 37, 42] success of similar Multi-level and Domain Decomposition solution algorithms for p and h - p version FE discretizations discussed in the previous section, we expect FVE to outperform FE in standard measures of computational efficiency.

Furthermore, from the standpoint of the *a priori* error analysis presented in the later chapters, we find that the accuracy of FVE approximations to the exact solution of general elliptic and parabolic problems will either match or surpass FE asymptotic convergence rates for h , p , and h - p versions of the respective methods.

2.5 Sobolev Spaces

In this section, we present the Sobolev space infrastructure for our FVE analysis. In the basic definitions and results, we have followed the well-known and comprehensive text of Ciarlet [11] unless otherwise noted. For a

more detailed and advanced discussion of Sobolev spaces, we refer the reader to the authoritative text of Adams [1] for mathematical theory and to the text of Kesavan [30] for applications to partial differential equations.

Throughout this section, $\Omega \subset \mathfrak{R}^d, d \in \{1, 2, 3\}$, denotes a *domain*: a bounded open connected set with Lipschitz-continuous boundary Γ . We also use the *multi-index notation* for denoting partial derivatives: let a multi-index α be denoted by a vector with non-negative integer components, $\alpha = (\alpha_1, \dots, \alpha_d) \in N^d$, equipped with the L^1 vector norm, $|\alpha| = \sum_{i=1}^d \alpha_i$. Taking $d = 2$ for definiteness and letting x and y denote the spatial coordinates of \mathfrak{R}^2 , we define $\partial^\alpha u$ as

$$\partial^\alpha u \equiv \partial_x^{\alpha_1} \partial_y^{\alpha_2} u, \quad (2.7)$$

so that $\partial^\alpha u$ represents the composition of the α_1 th and the α_2 th partial derivatives with respect to x and y .

2.5.1 Integer-Order Spaces

For any integer $k \geq 0$, and number $q \in [1, \infty]$, the *integer-order Sobolev space* $W^{k,q}(\Omega)$ consists of those functions $u \in L^q(\Omega)$ for which all partial derivatives $\partial^\alpha u$ (in sense of distributions) with $|\alpha| \leq k$ belong to the space $L^q(\Omega)$. The space $W^{k,q}(\Omega)$ is equipped with the norm:

$$\|u\|_{k,q,\Omega} \equiv \begin{cases} \left(\sum_{|\alpha| \leq k} \int_{\Omega} |\partial^\alpha u|^q d\mathbf{x} \right)^{1/q}, & q \in [1, \infty), \\ \max_{|\alpha| \leq k} \left\{ \text{ess sup}_{x \in \Omega} |\partial^\alpha u(x)| \right\}, & q = \infty. \end{cases} \quad (2.8)$$

The semi-norm $|\cdot|_{k,q,\Omega}$ is obtained by taking $|\alpha| = k$ in (2.8).

The space $W^{k,q}(\Omega)$ is a Banach space and a Hilbert space when $q = 2$.

In the latter case, we denote $W^{k,2}(\Omega)$ by $H^k(\Omega)$ and $\|\cdot\|_{k,2,\Omega}$ by $\|\cdot\|_{k,\Omega}$. If $u \in H^k(\Omega)$ for some $k \geq 0$, we say that u is *k-regular*. In our analysis, we will be using the $H^k(\Omega)$ spaces almost exclusively and will cite additional Sobolev space results (e.g., fractional-order spaces, trace theorem, embedding, etc.) in terms of this $q = 2$ case to simplify notation. These results have generalizations to $q \neq 2$ (see Adams [1] or Kesavan [30]), but they are not required in our forthcoming work.

2.5.2 Fractional-Order Spaces

In addition to the usual integer-order Sobolev spaces, we also employ *fractional-order Sobolev spaces* $H^S(\Omega)$, where $S \geq 0$ is a real number, as in Adams [1] and Kesavan [30]. In this discussion, S is decomposed into integer and fractional parts: i.e., $S = k + \theta$, where $k \geq 0$ is the integer part of S and $\theta \in [0, 1)$ is the fractional part of S . We define $H^S(\Omega)$ in terms of $H^k(\Omega)$ and $H^\theta(\Omega)$:

$$H^S(\Omega) \equiv \left\{ u \in H^k(\Omega) : \partial^\alpha u \in H^\theta(\Omega), \forall |\alpha| = k \right\}, \quad (2.9)$$

$$H^\theta(\Omega) \equiv \left\{ u \in L^2(\Omega) : \frac{|u(\mathbf{x}) - u(\mathbf{y})|}{|\mathbf{x} - \mathbf{y}|^{\theta+(d/2)}} \in L^2(\Omega \times \Omega) \right\}. \quad (2.10)$$

If $u \in H^S(\Omega)$ for some $S \geq 0$, we say that u is *S-regular*.

2.5.3 Sobolev Embedding

In our analysis, we approximate elements from fractional-order Sobolev spaces by C^0 piecewise-polynomial (Lagrange) interpolants. To ensure that interpolation is well-defined, we will only use the simplest embedding:

$$H^S(\Omega) \hookrightarrow C^0(\overline{\Omega}), \text{ if } S > d/2, \quad (2.11)$$

where $\hookrightarrow\hookrightarrow$ means that $H^S(\Omega)$ is contained in $C^0(\overline{\Omega})$ with a compact injection.

Referring to the embedding (2.11), we see that if u is $(1 + \epsilon)$ -regular for any $\epsilon > 0$, then u is represented by a C^0 function in \mathfrak{R}^2 . Therefore, u has point values and we can easily define C^0 Lagrange interpolation for u . Similarly, $3/2 + \epsilon$ regularity is required in three dimensions.

2.5.4 Trace Theorem

We will often employ the continuous and linear *trace map* of an element of a Sobolev space from an interior domain $\Omega \subset \mathfrak{R}^d, d \geq 2$, to a sub-dimensional boundary $\Gamma \subset \mathfrak{R}^{d-1}$ that is sufficiently smooth (see Kesavan [30] for details). The trace map, $\text{tr}(\cdot)$, maps $u \in H^{\Theta+1/2}(\Omega)$ to $\text{tr}(u) \in H^\Theta(\Gamma)$, where Θ is a positive, non-integer, real number (i.e., $\Theta \in \mathfrak{R}^+ \setminus \mathbb{N}$). When no confusion should arise, we simply write $\text{tr}(u)$ as u . The following *trace theorem* demonstrates the boundedness of the trace map: there exists a positive constant $C = C(\Omega)$ such that

$$\|u\|_{\Theta, \Gamma} \leq C \|u\|_{\Theta+1/2, \Omega}, \quad \Theta \in \mathfrak{R}^+ \setminus \mathbb{N}. \quad (2.12)$$

As a result of (2.12), we see that $u \in H^{S+1}(\Omega)$ for $S > 1/2$ ensures that $u \in H^1(\Gamma)$ and that the diffusive fluxes in (1.4) are bounded. With this regularity, C^0 interpolation is easily defined (up to three dimensions) as discussed in the embedding result above.

2.5.5 Poincaré Inequality

The Sobolev space $H_0^1(\Omega)$ is the closure of the space $\mathcal{D}(\Omega)$ —the space of infinitely differentiable functions with compact support in Ω —in the space $H^1(\Omega)$. More intuitively, $H_0^1(\Omega) \equiv \{u \in H^1(\Omega) : \text{tr}(u) = 0 \text{ on } \Gamma\}$. As a

result of the boundedness of Ω , we have the *Poincaré inequality*: there exists a positive constant $C(\Omega)$ such that

$$|u|_{0,\Omega} \leq C |u|_{1,\Omega}, \quad \forall u \in H_0^1(\Omega). \quad (2.13)$$

Therefore, $|\cdot|_{1,\Omega}$ is a norm over $H_0^1(\Omega)$ equivalent to $\|\cdot\|_{1,\Omega}$.

2.6 Finite Element Analysis

Here we present the rudiments of FE analysis to establish a reference point for analogous developments to follow for FVE. For simplicity, consider a steady diffusion equation with homogeneous Dirichlet boundary condition:

$$-\nabla \cdot (D\nabla u) = f, \quad x \in \Omega, \quad (2.14)$$

$$u|_{\partial\Omega} = 0, \quad (2.15)$$

where D is a continuously differentiable, bounded, non-degenerate ($0 < D_m \leq D \leq D_M < \infty$) diffusion coefficient and $\Omega \subset \mathfrak{R}^2$ is a domain.

A weak solution $u \in \mathcal{W} \equiv H_0^1(\Omega)$ to (2.14), corresponding to $f \in L^2(\Omega)$, satisfies:

$$A(u, w) = (f, w), \quad \forall w \in \mathcal{W}, \quad (2.16)$$

where

$$A(u, w) = \int_{\Omega} D\nabla u \cdot \nabla w \, d\mathbf{x}, \quad (2.17)$$

$$(f, w) = \int_{\Omega} f w \, d\mathbf{x}. \quad (2.18)$$

In our analysis, we assume additional regularity for u : i.e., u lies in the (fractional-order) Sobolev space $\mathcal{W}^+ \equiv H^{S+1}(\Omega) \cap \mathcal{W}$, where $S \in (1/2, p]$ and $p \geq 1$. The interval $(1, p+1]$ for $S+1$ is the full range of admissible regularities for u in FE analysis for a general unspecified problem domain. (Technically, the lower bound on S can be reduced to 0 in two dimensions; however, 1/2 is required for our results to hold in three dimensions.) But for any specific problem domain, the value of $S+1$ in that range is precisely determined: $S = S(\partial\Omega)$ is a function of the smoothness of the boundary $\partial\Omega$ and the geometry of any corners in non-smooth domains (see Grisvard [19] for details). For example, if

Ω is an L-shaped domain, $S + 1 = 5/3 - \epsilon \forall \epsilon > 0$. To justify the upper end of the regularity scale ($S + 1 > 3$) associated with high-order polynomial trial spaces ($p > 2$), we assume that curved elements are being used to approximate the boundaries of smooth domains (see Ciarlet [11] for details).

In the FE analysis presented here and in the FVE analysis of the following chapters, we invoke fractional-order spaces to demonstrate error estimates across the fullest range of admissible regularities to encompass any particular problem domain.

Also, the numerical solution u^h of (2.14) (see (2.19) below) will lie in $\mathcal{W}^h \equiv \mathcal{S}_0^p(\mathcal{T}^h) \subset \mathcal{W}$ —the space of C^0 functions that are polynomials of order $\leq p$ on each element T of the triangulation \mathcal{T}^h and vanish on $\partial\Omega$.

In direct analogy with (2.16), the (finite element) numerical solution $u^h \in \mathcal{W}^h$ to (2.14), corresponding to $f^h \in \mathcal{S}^p(\mathcal{T}^h)$, satisfies:

$$A(u^h, w) = (f^h, w), \quad \forall w \in \mathcal{W}^h. \quad (2.19)$$

Here, f^h is the $L^2(\Omega)$ projection of f into $\mathcal{S}^p(\mathcal{T}^h)$:

$$(f - f^h, w) = 0, \quad \forall w \in \mathcal{S}^p(\mathcal{T}^h). \quad (2.20)$$

The key components of finite element error analysis are boundedness, approximation theory, and ellipticity results; we introduce and discuss each below.

2.6.1 Boundedness

For the bilinear form $A(\cdot, \cdot)$ of (2.17)—hereafter called the “ A -form”—consider the following upper bound (or *continuity*) condition: there exists a

positive constant M such that

$$A(u, w) \leq M |u|_{1,\Omega} |w|_{1,\Omega}, \quad \forall u, w \in \mathcal{W}. \quad (2.21)$$

Such a condition is usually assumed to obtain convergence of the numerical method as seen below.

2.6.2 Approximation Theory

At the heart of FE analysis are error estimates for approximation of elements from Sobolev spaces by elements of piecewise polynomial FE trial spaces. For the $H^k(\Omega)$ norms ($k \in \{0, 1\}$) that we shall employ, we cite standard approximation theory results for polynomial interpolation (see Babuška *et al.* [2, 3], Jensen and Suri [27], and Dupont and Scott [16]): there exists a constant C such that

$$\|\eta_p(u)\|_{k,\Omega} \leq C (h/p)^{S+1-k} \|u\|_{S+1,\Omega}, \quad S \in (1/2, p], \quad (2.22)$$

where $\eta_p(u) \equiv \Pi_p^h u - u$ is the interpolation error for u in the polynomial space $\mathcal{S}^p(\mathcal{T}^h)$. As discussed above for an L-shaped domain, (2.22) holds for $S = 2/3 - \epsilon$. In three dimensions, the lower bound on S in (2.22) is necessitated by the Sobolev embedding result (2.11), $3/2 + \epsilon$ regularity is required to embed into $C^0(\overline{\Omega})$ and easily define interpolation.

In accordance with the definitions of Section 2.3, the term $(h/p)^S$ in (2.22) is understood as $C(p)h^S$ for h -version methods and as $C(h)p^{-S}$ for p -version methods. In the remainder of the text, C will represent a (generic) positive constant.

2.6.3 Ellipticity

Consider the following *ellipticity* (or lower bound) condition: there exists a positive constant β such that

$$A(u, u) \geq \beta |u|_{1,\Omega}^2, \quad \forall u \in \mathcal{W}. \quad (2.23)$$

This condition is perhaps the most important technical component of FE error analysis: it implies that the A -form is positive definite. Once ellipticity is established, FE analysis reduces to the approximation theory of the previous subsection.

2.6.4 Error Analysis

In the standard FE error analysis to follow, we apply the results (2.23) and (2.21)–(2.22); for (2.14), (2.23) and (2.21) hold with $\beta = D_m$ and $M = D_M$, respectively.

2.6.4.1 Error Splitting

Let $e \equiv u - u^h \in \mathcal{W}$ be the error in the FE approximation to (2.14). For theoretical purposes, define $\zeta \equiv \Pi_p^h e$ as an interpolation of the error from \mathcal{W} into \mathcal{W}^h , so that $\zeta \equiv \Pi_p^h u - u^h \in \mathcal{W}^h$. We refer to ζ as the *representation error* in \mathcal{W}^h , i.e., the difference between interpolated and computed approximations in \mathcal{W}^h of the exact solution u . Then, we employ the *error splitting* $e = \zeta - \eta_p$, in which the FE error is split into its representation error and interpolation error components: $|u - u^h|_{1,\Omega} \leq |\zeta|_{1,\Omega} + |\eta_p|_{1,\Omega}$. With the estimate of $|\eta_p|_{1,\Omega}$ in (2.22), estimating $|e|_{1,\Omega}$ is reduced to estimating $|\zeta|_{1,\Omega}$. Below, we develop an estimate for $|\zeta|_{1,\Omega}$ in terms of $|\eta_p|_{1,\Omega}$.

2.6.4.2 Zero Property

For all $w \in \mathcal{W}^h$, the exact solution satisfies (2.16), the numerical solution satisfies (2.19), and the source term satisfies (2.20). Subtracting, we have the *zero property* or *orthogonality* of the error,

$$A(u - u^h, w) = 0, \quad \forall w \in \mathcal{W}^h, \quad (2.24)$$

which in our context is more appropriately expressed as

$$A(\zeta, w) = A(\eta_p, w), \quad \forall w \in \mathcal{W}^h. \quad (2.25)$$

2.6.4.3 Error Estimate

Since the representation error, ζ , is in \mathcal{W}^h , we take $w = \zeta$ in (2.25) to obtain

$$\begin{aligned} \beta |\zeta|_{1,\Omega}^2 &\leq A(\zeta, \zeta), \\ &= A(\eta_p, \zeta), \\ &\leq M |\eta_p|_{1,\Omega} |\zeta|_{1,\Omega}. \end{aligned} \quad (2.26)$$

The three lines of (2.26) use (2.23), (2.25), and (2.21), respectively. By (2.22), (2.26) implies the following H^1 optimal-order *error estimate* for the finite element method:

$$|u - u^h|_{1,\Omega} \leq C (h/p)^S \|u\|_{S+1,\Omega}, \quad S \in (1/2, p]. \quad (2.27)$$

In the FVE analysis to follow, we obtain error estimates that match or surpass the p th-order estimate (2.27).

2.7 Summary

In this chapter, we presented some theoretical preliminaries. In Sections 2.1 and 2.2, we defined and discussed the computational meshes (i.e., the FE triangulation \mathcal{T}^h , the FVE primal mesh $\mathcal{T}^{h/p}$, and the FVE dual mesh $\mathcal{V}^{h/p}$) that are fundamental to the implementation and analysis of FVE. In Sections 2.3 and 2.4, we defined h , p , and h - p version FE and FVE and discussed their relative merits. In Section 2.5, we introduced standard definitions and results for Sobolev spaces that are fundamental to FE and FVE analysis. In Section 2.6, we briefly outlined an FE analysis for the steady diffusion equation to set a reference point for an FVE analysis in the next two chapters.

3. Analysis for Diffusion Equations

In this chapter, we present an h , p , and h - p version FVE analysis for diffusion equations. In Section 3.1, we provide some preliminary definitions and notation that facilitates the transition from FE to FVE analysis. In Section 3.2, we motivate our FVE analysis and its relation to previous FE work. In Section 3.3, we define some discrete norm counterparts to the continuous L^2 and H^1 Sobolev norms of Section 2.5 and demonstrate a discrete Poincaré inequality. As in the FE analysis of Section 2.6, these tools are crucial to FVE analysis. In Section 3.4, we determine an upper bound or continuity result in terms of an FVE diffusion functional and our discrete H^1 semi-norm for the FVE bilinear form that is used to discretize the model diffusion equation discussed in Section 2.6. In Section 3.5, we state and outline applications of the linear Bramble-Hilbert lemma which is at the heart of our FVE analysis. In Section 3.6, we use the Bramble-Hilbert lemma to derive error estimates for the FVE approximation of elements from fractional-order Sobolev spaces by elements from the piecewise polynomial trial spaces employed by the h , p , and h - p version FVE methods: in terms of our FVE diffusion functional, $H^1(\Omega)$ -equivalent optimal-order approximation results are derived for p th-order trial spaces in Section 3.6.1; robust and practical superconvergence results for linear trial spaces on a “symmetric” volumization are derived in two and three dimensions in Section 3.6.2. In Section 3.7, we derive an “asymptotic” discrete ellipticity result for the FVE diffusion bilinear form in terms of our discrete

H^1 semi-norm—aided by the approximation theory from the previous section. Finally in Section 3.8, we construct an FVE analysis for diffusion equations with the components developed in the preceding sections: optimal-order and superconvergence results are obtained for h , p , and h - p version FVE in our discrete H^1 semi-norm.

3.1 Preliminaries

We employ the same continuous and discrete space notations, definitions, and structures as in the previous section for the FVE continuous and discrete trial spaces that represent the solution to (2.14): $u \in \mathcal{W} \equiv H_0^1(\Omega)$ and $u^h \in \mathcal{W}^h \equiv \mathcal{S}_0^p(\mathcal{T}^h)$, for $p \geq 1$. Similarly in FVE analysis, we assume additional regularity for u : u lies in the (fractional-order) Sobolev space $\mathcal{W}^+ \equiv H^{S+1}(\Omega) \cap \mathcal{W}$, where $S \in (1/2, p]$. The $1/2$ lower limit on S ensures: a bounded diffusive flux, due to the Sobolev trace theorem (2.12) as discussed below; and well-defined interpolations in up to three spatial dimensions, due to the Sobolev embedding (2.11). As stated at the start of Section 2.6, $(3/2, p + 1]$ is the full range of admissible regularities in our analysis for an unspecified problem domain. For any specified Ω , the value of S in that range is a function of the smoothness of $\partial\Omega$ [19]: e.g., on an L-shaped domain, $S = 2/3 - \epsilon \forall \epsilon > 0$. In the remainder of the text, ϵ will represent a (generic) infinitesimal positive constant. Also, to justify the upper end of the regularity scale ($S + 1 > 3$) associated with high-order polynomial trial spaces ($p > 2$), we assume that curved elements are being used to approximate the boundaries of smooth domains (see Ciarlet [11] for details).

In addition, let \mathcal{X}^h denote the FVE test space $P_0^0(\mathcal{V}^{h/p})$ —the polynomials that are constant on each volume V of $\mathcal{V}^{h/p}$ and vanish on $\partial\Omega$. Thus, test functions corresponding to volumes that intersect $\partial\Omega$ are identically zero. Also, let χ_V denote the characteristic function associated with volume $V \subseteq \overline{\Omega}$ and let \mathcal{V} denote the collection of all such V .

Recalling the integral equation (1.4), a weak solution u of (2.14) satisfies the following problem: given $f \in L^2(\Omega)$, find $u \in \mathcal{W}^+$ such that

$$B(u, \chi_V) = (f, \chi_V), \quad \forall V \in \mathcal{V}, \quad (3.1)$$

where

$$B(u, \chi_V) = - \int_{\partial V} D\nabla u \cdot \mathbf{n} \, dS, \quad (3.2)$$

$$(f, \chi_V) = \int_V f \, d\mathbf{x}. \quad (3.3)$$

FVE poses the weak form (3.1) on the volumization $\mathcal{V}^{h/p}$ —a finite subset of \mathcal{V} with non-overlapping volumes—and employs standard FE representations of u and f . Thus, FVE replaces (3.1) with: given $f^h \in \mathcal{S}^p(\mathcal{T}^h)$, find $u^h \in \mathcal{W}^h$ such that

$$B(u^h, \chi_V) = (f^h, \chi_V), \quad \forall V \in \mathcal{V}^{h/p}. \quad (3.4)$$

Here, f^h satisfies a non-standard $L^2(\Omega)$ projection relationship with f :

$$(f - f^h, \chi_V) = 0, \quad \forall V \in \mathcal{V}^{h/p}, \quad (3.5)$$

which ensures that the numerical source term conserves mass on each volume.

Since $\mathcal{V}^{h/p}$ partitions Ω and the volumes in $\mathcal{V}^{h/p}$ are related to the finite element triangulation \mathcal{T}^h through $\mathcal{T}^{h/p}$ as discussed in Section 2.2, (3.4)

is more amenable than (3.1) to an FE-style analysis. Below, we point out similarities and differences between FVE and FE from the perspective of an FE-style framework for analysis.

3.2 Motivation

To motivate our FVE analysis and its relation to previous FE work, reconsider (2.14)–(2.15) and the FE analysis of Section 2.6. We saw that the crucial ellipticity result (2.23) for the A -form and boundedness result (2.21) are stated in terms of the continuum space \mathcal{W} , but the results also held for the discrete space \mathcal{W}^h (i.e., $\mathcal{W}^h \subset \mathcal{W}$). Superficially, it could appear that the focus was on the continuum problem (2.16) rather than on the FE discrete problem (2.19). However, in Section 2.6.4, the error analysis emphasized results for the discrete FE problem (2.19) and the associated discrete spaces. Recall that the exact weak solution u was invoked only in the zero property (2.25); furthermore, recall that the key observation in (2.25) was that u satisfied the FE discrete equation (2.19). Therefore, to begin an FE-style analysis for FVE, we need to generate results analogous to those of Sections 2.6.1, 2.6.2 and 2.6.3 for the FVE discrete problem (3.4) rather than for (3.1).

The continuum problems (2.16) and (3.1) are different in physical character (minimization vs. conservation) as well as in mathematical content and these differences discourage a single, unifying framework for their analysis. For the discrete problems (2.19) and (3.4), the situation is more favorable for a unified analysis. We now manipulate (3.4) to be more reminiscent of (2.19).

If $w \in \mathcal{X}^h$, w has a volume-wise constant representation:

$$w \equiv \sum_{i \in I} w_i \chi_i, \quad (3.6)$$

where $\chi_i = \chi_{V_i}$ and $\mathcal{V}^{h/p} = \{V_i : i \in I\}$. In terms of w , the FVE solution $u^h \in \mathcal{W}^h$ of (3.4) also satisfies:

$$B(u^h, w) = (f^h, w), \quad \forall w \in \mathcal{X}^h, \quad (3.7)$$

where

$$B(u^h, w) = \sum_{i \in I} B(u^h, w_i \chi_i), \quad (3.8)$$

$$(f^h, w) = \sum_{i \in I} (f^h, w_i \chi_i). \quad (3.9)$$

That is, if u^h satisfies the each of the local equations of (3.4), then, by (3.7), u^h also satisfies any linear combination of those equations. Notice that (3.8) and (3.9) are now global in extent, like their FE counterparts (2.17) and (2.18), and (3.7) with $w \equiv 1$ represents a global conservation law on the union or aggregate of all volumes in $\mathcal{V}^{h/p}$. Like the FE A -form, we refer to (3.8) as the B -form.

Similarly, the $L^2(\Omega)$ projection relationship (3.5) for $f^h \in \mathcal{S}^p(\mathcal{T}^h)$ is transformed into

$$(f - f^h, w) = 0, \quad \forall w \in \mathcal{X}^h. \quad (3.10)$$

Now we establish some additional theoretical infrastructure to support a FE-style analysis of FVE based on (3.7).

3.3 Discrete Norms

In this section, we define some discrete norm counterparts to the continuous integer-ordered Sobolev norms of Section 2.5 and demonstrate a discrete Poincaré inequality. As in FE analysis of Section 2.6, these tools are crucial to FVE analysis.

In order to distinguish between continuous and discrete norms in our definitions, we replace the physical domain Ω that was employed in the definition of the continuous Sobolev norms with the set of computational mesh points ω (i.e., the vertices of the triangulation $\mathcal{T}^{h/p}$) contained in $\bar{\Omega}$ in our definition of discrete Sobolev norms. Recall that the continuous norms were defined in terms of integrated functions over Ω ; however, our discrete norms will be defined in terms of function point values in ω . Therefore, only Sobolev spaces that can be embedded into $C^0(\bar{\Omega})$, like $H^S(\Omega)$ for $S > 1/2$, are admissible in a discrete norm analysis.

In particular, we will restrict our attention to certain polynomial subspaces of $C^0(\bar{\Omega})$ when defining certain discrete Sobolev norms and semi-norms: $\mathcal{S}^0(\mathcal{V}^{h/p})$ for the $L^2(\omega)$ norm and $\mathcal{S}^1(\mathcal{T}^{h/p})$ for the $H^1(\omega)$ semi-norm. This choice of polynomial spaces will emphasize and exploit the relationships between the FVE dual and primal meshes, which are at the heart of any FVE implementation, in our FVE analysis. Thus, discrete norms will provide for a unified analysis of both practical and theoretical questions in our study of FVE: attention will be shifted from estimating the error over the whole domain Ω to estimating the error over only the subset of points ω where our numerical method is approximating the exact solution to the original partial differential

equation (1.1). Starting with Section 3.6, we find that FVE discrete norm analysis reveals nodal superconvergence in multiple dimensions as well as the usual optimal-order convergence rates and is superior to the FE continuous norm analysis of Section 2.6.

3.3.1 Discrete L^2 Norm

First, we define a discrete variant of the standard continuous $L^2(\Omega)$ norm—referred to as a $L^2(\omega)$ norm.

Definition 3.1 The $L^2(\omega)$ norm is given by

$$|u|_{0,\omega} \equiv \left(\sum_{i \in I} u_i^2 |V_i| \right)^{1/2}, \quad (3.11)$$

where

- $\omega \equiv \{\mathbf{x}_i : i \in I\}$ is the set of primal mesh points;
- I is the primal mesh point indexing set;
- u_i is the value of u at the primal mesh point \mathbf{x}_i ;
- $V_i \in \mathcal{V}^{h/p}$ is the volume that corresponds to \mathbf{x}_i ;
- $|V_i|$ is the standard Lebesgue volume measure of V_i .

Clearly, the $L^2(\omega)$ norm is equivalent (cf. Bank and Rose [5]) to the $L^2(\Omega)$ norm restricted to $\mathcal{S}^0(\mathcal{V}^{h/p})$ —the space of piecewise constant functions on the FVE dual mesh $\mathcal{V}^{h/p}$:

$$|u|_{0,\omega} \equiv |\Pi_0^{h/p} u|_{0,\Omega}, \quad (3.12)$$

where $\Pi_0^{h/p} u$ denotes interpolation into $\mathcal{S}^0(\mathcal{V}^{h/p})$. Then, $L^2(\omega)$ is the space of functions (generating primal mesh point sequences) that are bounded in the $L^2(\omega)$ norm.

3.3.2 Discrete H^1 Norm

As in the FE analysis of Section 2.6, the H^1 semi-norm plays a pivotal role in the course of FVE analysis. By analogy, we define a discrete variant of the standard continuous $H^1(\Omega)$ semi-norm—referred to as an $H^1(\omega)$ semi-norm.

Definition 3.2 The $H^1(\omega)$ semi-norm is given by

$$|u|_{1,\omega} \equiv \left(\sum_{\{i,j\} \in W} (u_i - u_j)^2 \frac{|\gamma_{ij}|}{|X_{ij}|} \right)^{1/2}, \quad (3.13)$$

where

- $W \equiv \{\{i, j\} : i \in I, j \in W_i \text{ with } |\gamma_{ij}| > 0\}$ is a set of unordered index pairs corresponding to the edges of each element $T_p \in \mathcal{T}^{h/p}$;
- $W_i \equiv \{j : \mathbf{x}_j \in \mathcal{N}(i), j \in I\}$ is a set of indices for \mathbf{x}_i 's nearest neighbor mesh points;
- $\mathcal{N}(i)$ is a set of \mathbf{x}_i 's nearest neighbor mesh points in ω ;
- $|\gamma_{ij}|$ is the standard Lebesgue surface measure of the volume interface $\gamma_{ij} \equiv \overline{V}_i \cap \overline{V}_j$; in 1-D, replace $|\gamma_{ij}|$ with 1;
- $|X_{ij}|$ is the Euclidean distance between \mathbf{x}_i and \mathbf{x}_j .

Despite outward appearances, the $H^1(\omega)$ semi-norm is equivalent (cf. Bank and Rose [5]) to the standard $H^1(\Omega)$ semi-norm restricted to $\mathcal{S}^1(\mathcal{T}^{h/p})$ —the space of C^0 piecewise linear polynomial functions on the FVE primal mesh $\mathcal{T}^{h/p}$:

$$|u|_{1,\omega} \equiv |\Pi_1^{h/p} u|_{1,\Omega}, \quad (3.14)$$

where $\Pi_1^{h/p} u$ denotes interpolation into $\mathcal{S}^1(\mathcal{T}^{h/p})$. Then, $H^1(\omega)$ is the space of functions that are bounded in the $H^1(\omega)$ norm:

$$\|u\|_{0,\omega} \equiv \left(|u|_{0,\omega}^2 + |u|_{1,\omega}^2 \right)^{1/2}. \quad (3.15)$$

To relate $|\cdot|_{0,\omega}$ and $|\cdot|_{1,\omega}$ more directly, we demonstrate a discrete Poincaré inequality.

3.3.3 Discrete Poincaré Inequality

Following the continuous case introduced in Section 2.5, let γ denote the set of primal mesh points that lie on the domain boundary Γ . By analogy, define $H_0^1(\omega)$ as the set $\{u \in H^1(\omega) : u = 0 \text{ on } \gamma\}$. Then the boundedness of Ω translates to a finite number of mesh points in ω and to the following result.

Lemma 3.1 There exists a positive constant C such that

$$|u|_{0,\omega} \leq C |u|_{1,\omega}, \quad \forall u \in H_0^1(\omega). \quad (3.16)$$

Therefore, $|\cdot|_{1,\omega}$ is a norm over $H_0^1(\omega)$ equivalent to $\|\cdot\|_{1,\omega}$.

Proof: Lemma 2 of Bank and Rose [5] demonstrates that $|\Pi_0^{h/p} u|_{0,\Omega} \leq C |\Pi_1^{h/p} u|_{0,\Omega}$. Since $\Pi_1^{h/p} u \in \mathcal{S}_0^1(\mathcal{T}^{h/p}) \subset H_0^1(\Omega)$, the continuous Poincaré inequality (2.13) yields $|\Pi_1^{h/p} u|_{0,\Omega} \leq C |\Pi_1^{h/p} u|_{1,\Omega}$. Therefore, application of the equivalences (3.12) and (3.14) gives us the desired result (3.16). \square

As in FE error analysis, the key components of FVE analysis are boundedness, approximation theory, and ellipticity results: we demonstrate and discuss each result below. All follow the basic FE-style framework introduced in Sections 2.6.1–2.6.3 after discrete norm modifications.

3.4 Boundedness

To determine an upper bound for the FVE B -form (3.8), we are guided by the corresponding development for the FE A -form (2.17). In Section 2.6.1, the upper bound (2.21) for the A -form is given in terms of a product of standard integer-order Sobolev norms (or sub-linear functionals). In the error analysis of Section 2.6.4, these norms are applied to a continuous trial function and a discrete test function. Since this upper bound is used in conjunction with the ellipticity result (2.23), the only theoretical constraint on the bound is that the norm applied to the test function must match (or can be made to match via an auxiliary result like the Poincaré inequality) the norm used in the ellipticity result—i.e., $|\cdot|_{1,\Omega}$ for FE and $|\cdot|_{1,\omega}$ for FVE.

In FE analysis, the $H_0^1(\Omega)$ norm constraint on the test function lead to the choice of the $H_0^1(\Omega)$ norm for the trial functions. However, in FVE analysis, the $H_0^1(\omega)$ norm constraint on the test function leads to the choice of a problem-dependent, bounded functional for the trial functions that corresponds to the FVE treatment of diffusion in (1.4). In Section 4.2, we will develop functionals that correspond to FVE treatments of reaction and advection in (1.4). Practically speaking, we want this diffusion functional (and, later on, reaction and advection functionals) to admit optimal-order or even superconvergent approximation theorems, since the continuous trial function in our analysis is the interpolation error for the exact solution.

In terms of the admissible trial space, $\mathcal{H} \equiv \mathcal{W}^+ \oplus \mathcal{W}^h$, and test space \mathcal{X}^h for FVE, we demonstrate an boundedness result for the B -form (3.8) (cf. (2.21) of Section 2.6.1).

Lemma 3.2 (Diffusion bound)

$$|B(u, w)| \leq \mathcal{D}(u) |w|_{1, \omega}, \quad \forall u \in \mathcal{H}, w \in \mathcal{X}^h. \quad (3.17)$$

$\mathcal{D}(\cdot)$ is a bounded functional:

$$\mathcal{D}(u) \equiv \left(\sum_{\{i,j\} \in W} d_{ij}^2(u) \frac{|X_{ij}|}{|\gamma_{ij}|} \right)^{1/2}; \quad (3.18)$$

$d_{ij}(\cdot)$ is a bounded linear functional arising from the FVE treatment of diffusion:

$$d_{ij}(u) \equiv - \int_{\gamma_{ij}} D \nabla u \cdot \mathbf{n}_{ij} dS, \quad (3.19)$$

where \mathbf{n}_{ij} is the unit normal vector pointing outward from V_i into V_j .

Proof: By the Sobolev trace theorem, $u \in H^{S+1}(\Omega)$ for $S > 1/2$ ensures that (3.19), hence (3.18), is bounded. From the definition (3.8) of the B -form and the fact that

$$- \int_{\partial V_i} D \nabla u \cdot \mathbf{n}_i dS \equiv \sum_{j \in W_i} d_{ij}(u), \quad (3.20)$$

we see that

$$\begin{aligned} |B(u, w)| &= \left| \sum_{i \in I} w_i \sum_{j \in W_i} d_{ij}(u) \right|, \\ &= \left| \sum_{\{i,j\} \in W} d_{ij}(u) (w_i - w_j) \right|, \\ &\leq \left(\sum_{\{i,j\} \in W} d_{ij}^2(u) \frac{|X_{ij}|}{|\gamma_{ij}|} \right)^{1/2} \left(\sum_{\{i,j\} \in W} (w_i - w_j)^2 \frac{|\gamma_{ij}|}{|X_{ij}|} \right)^{1/2}, \end{aligned} \quad (3.21)$$

where we have used the fact that $d_{ij}(u) = -d_{ji}(u)$ (i.e., $\mathbf{n}_{ij} = -\mathbf{n}_{ji}$) in the second line of (3.21). \square

With an upper bound for the B -form in terms of the non-standard, problem-dependent, bounded linear functional (3.19), we need an approximation theorem: the theoretical basis for this result is introduced and discussed below.

3.5 Bramble-Hilbert Lemma

In this section, we state and outline applications of the linear Bramble-Hilbert lemma. This lemma was originally developed by Bramble and Hilbert [7] for use in approximation theory. Later, the lemma was modified for use in FE theory by Ciarlet [11] and applied in linear and bilinear forms. In the standard FE analysis presented in Section 2.6, this lemma only plays a supporting role—a bilinear Bramble-Hilbert lemma [11] is employed to analyze the effects of numerical quadrature. However, in modern FV theory the lemma plays a central role. In fact, the particular form of the Bramble-Hilbert lemma stated here was used in the pioneering FV analysis of Samarskii, Lazarov, and Makarov [46]. Similarly, the Bramble-Hilbert lemma plays a central role in our FVE error analysis: the lemma is used to demonstrate optimal-order and superconvergence error for bounded linear functionals that arise in the implementation and analysis of the finite volume element method.

First, we introduce some additional infrastructure for bounded linear functionals. If f is a bounded linear functional for elements of $H^S(\Omega)$ with $S \geq 0$, f is in the *dual space* of $H^S(\Omega)$ denoted by $H^{-S}(\Omega)$. The norm of f in $H^{-S}(\Omega)$ is

$$\|f\|_{-S,\Omega} \equiv \sup_{u \in H^S(\Omega)} \frac{|(f, u)|}{\|u\|_{S,\Omega}}, \quad (3.22)$$

where (\cdot, \cdot) denotes the duality pairing between f and u in terms of the standard $L^2(\Omega)$ inner product.

Now we introduce the linear Bramble-Hilbert lemma.

Lemma 3.3 (Bramble-Hilbert) Let Ω be a domain in \mathfrak{R}^d , $k \geq 0$ an integer, $\theta \in (0, 1]$, and f a bounded linear functional on the space $H^{k+\theta}(\Omega)$

that vanishes for all polynomials of degree $\leq k$ on Ω :

$$f(\psi) = 0, \quad \forall \psi \in \mathcal{S}^k(\Omega). \quad (3.23)$$

Then there exists a positive constant $C = C(\Omega)$ such that

$$|f(u)| \leq C \|f\|_{-(k+\theta), \Omega} |u|_{k+\theta, \Omega}, \quad \forall u \in H^{k+\theta}(\Omega). \quad (3.24)$$

3.5.1 Applications

Now that we have stated the linear Bramble-Hilbert lemma, we demonstrate its application for some functionals arising from standard h -version finite element approximations. For simplicity, we work in two dimensions and assume that the regularity of u is in the interval $(1, p + 1]$.

Let $T \in \mathcal{T}^h$ be an element (i.e., a triangular region in \mathbb{R}^2) in a FE triangulation of Ω , u a function defined on T , and $\Pi_1 u$ denote the linear interpolant of u on T . Then, $\eta = \Pi_1 u - u$ is the linear interpolation error on T . By the Bramble-Hilbert lemma, we have the following error estimate:

Example 3.1 (Linear interpolation)

$$|f(u)| \equiv \left| \int_T \eta \, d\mathbf{x} \right| \leq C h_T^{S+1} |T|^{1/2} |u|_{S+1, T}, \quad S \in (0, 1], \quad (3.25)$$

where h_T is the diameter of T .

Proof: We map the computational element T to a reference element \hat{T} by an affine map, $\mathbf{x} = \mathbf{M}\hat{\mathbf{x}} + \mathbf{c}$, and we transform our functional:

$$|f(u)| = \frac{|T|}{|\hat{T}|} |\hat{f}(\hat{u})|,$$

where

$$\hat{f}(\hat{u}) \equiv \int_{\hat{T}} \hat{\eta} \, d\hat{\mathbf{x}},$$

and $\eta(\mathbf{x}) = \hat{\eta}(\hat{\mathbf{x}})$ by the affine map. Since linear interpolation is exact for linear polynomials, the Bramble-Hilbert lemma says that

$$|\hat{f}(\hat{u})| \leq C(\hat{T}) \|\hat{f}\|_{-(S+1), \hat{T}} |\hat{u}|_{S+1, \hat{T}}, \quad S \in (0, 1].$$

Next, we note that $\|\hat{f}\|_{-(S+1), \hat{T}} \leq C|\hat{T}|^{1/2}$ and following Ciarlet [11] we transform back to T as in,

$$|\hat{u}|_{S+1, \hat{T}} \leq C \frac{\|\mathbf{M}\|^{S+1}}{(\det \mathbf{M})^{1/2}} |u|_{S+1, T}, \quad (3.26)$$

where $\|\mathbf{M}\| \leq C h_T / \rho_{\hat{T}}$ and $\rho_{\hat{T}}$ is the diameter of the circle inscribed in \hat{T} , and $\det \mathbf{M} = |T|/|\hat{T}|$. Since \hat{T} is a reference or fixed element, \hat{T} is chosen so that $\rho_{\hat{T}} = O(1)$. After transforming back to T and simplifying terms, we obtain the stated result. \square

In more generality, we obtain h -version results that are equivalent to optimal-order L^2 results for approximation by polynomials of degree $p \geq 1$:

Example 3.2 (p th-order interpolation)

$$\left| \int_T (\Pi_p u - u) d\mathbf{x} \right| \leq C h_T^{S+1} |T|^{1/2} |u|_{S+1, T}, \quad S \in (0, p]. \quad (3.27)$$

Proof: Since p th-order interpolation is exact, $\Pi_p u - u = 0$, when u is polynomial of degree $k \leq p$ on T , the Bramble-Hilbert lemma holds when the regularity parameter S for u satisfies $1 < S + 1 \leq p + 1$. We follow the affine mapping procedure discussed in the previous example and apply Bramble-Hilbert on the reference element. After the inverse mapping procedure to the original element, we gain $S + 1$ powers of h_T in the final result (3.27). \square

Finally, the estimates for approximation of first-order derivatives of u by polynomials of degree $p \geq 1$ are reduced by a power of h_T (cf. Ciarlet [11], Theorems 15.1-2).

Example 3.3 (First-order derivatives)

$$\left| \int_T \partial_x (\Pi_p u - u) d\mathbf{x} \right| \leq C h_T^S |T|^{1/2} |u|_{S+1,T}, \quad S \in (0, p]; \quad (3.28)$$

Proof: We follow the proof of the previous example with only one modification. When applying the affine mapping argument, note that we have the scaling relationship $\hat{x} = C \rho_T^{-1} x$. Then after transforming the first-order derivative by the chain rule, we find that

$$\partial_x u = C \rho_T^{-1} \partial_{\hat{x}} \hat{u}. \quad (3.29)$$

After applying the shape regularity inequality (2.1), the estimate of the previous example is reduced by a power h_T . \square

The h -version estimate (3.28) is equivalent to optimal-order H^1 results for approximation by polynomials of degree $p \geq 1$. Since FE uses optimal-order H^1 results to obtain error estimates as in (2.22), use of the Bramble-Hilbert lemma does not improve the results of standard FE analysis. However, FVE uses different functionals than those stated here (i.e., functionals based on volume elements and not on finite elements) and the Bramble-Hilbert lemma can exploit special volumes symmetries to add a power of h to the ($p = 1$) approximation result (3.28) for derivatives.

3.6 Approximation Theory

As in FE analysis, error estimates for approximation of elements from fractional-order Sobolev spaces by elements from piecewise polynomial FE trial spaces are at the heart of FVE analysis. By applying the Bramble-Hilbert lemma to the bounded linear functional (3.19), we demonstrate approximation theorems for the diffusion functional $\mathcal{D}(\cdot)$. Our baseline criterion is that our results be equivalent to $H^1(\Omega)$ optimal-order estimates for polynomial interpolation (as stated in (2.22) of Section 2.6.2).

To justify the upper end of the regularity scale ($S + 1 > 3$) associated with high-order polynomial trial spaces ($p > 2$), we assume that curved (isoparametric-equivalent) elements are being used to approximate the curved boundaries of smooth domains. However, we ensure the invertibility of the isoparametric mappings to define these curved elements by requiring a boundary discretization sufficiently fine to bound or approximate the isoparametric map with an affine map (see Ciarlet [11] or Johnson [28] for details). Thus, it is sufficient to work with the affine mappings that generate affine-equivalent families of polygonal elements. Hence, all of our approximation theory results in this dissertation are demonstrated for polygonal finite elements under affine mappings without loss of generality.

3.6.1 Optimal-order Approximation

First, we demonstrate an ($H^1(\Omega)$ -equivalent) optimal-order result for p th-order interpolation when the triangulation \mathcal{T}^h and volumization $\mathcal{V}^{h/p}$ are regular (defined in Sections 2.1.1 and 2.2.3)—cf. Theorem 17.1 of Ciarlet [11].

According to the definitions of Section 2.3, the term $(h/p)^S$ in (3.30) is understood as $C(p)h^S$ for h -version FVE and as $C(h)p^{-S}$ for p -version FVE.

Theorem 3.1 Let $\eta_p(u) = \Pi_p^h u - u$ be the interpolation error for $u \in \mathcal{W}^+$ into $\mathcal{S}^p(\mathcal{T}^h)$. Then we have the following estimate:

$$\mathcal{D}(\eta_p(u)) \leq C (h/p)^S \|u\|_{S+1, \Omega}, \quad S \in (1/2, p]. \quad (3.30)$$

Proof: As in Cai [10], define a “macro-element” E_{ij} as the region formed by joining the endpoints of γ_{ij} and X_{ij} (see Figure 2.6). As noted in Section 2.2.3, E_{ij} is the union of two sub-elements (with X_{ij} as a common edge) in an auxiliary triangulation $\widetilde{\mathcal{T}}^{h/p}$ which is assumed to be shape regular. Thus, $E_{ij} = T_{ij}^1 \cup T_{ij}^2$, where each sub-element is contained in a single element of $\mathcal{T}^{h/p}$.

Recalling the nesting of the triangulations $\widetilde{\mathcal{T}}^{h/p} \subseteq \mathcal{T}^{h/p}$ in (2.6), let T_p^k be the element of $\mathcal{T}^{h/p}$ that contains T_{ij}^k for $k \in \{1, 2\}$. In our analysis, polynomial functions and their interpolants are defined on T_p^k and then restricted to T_{ij}^k .

Then we work on each sub-element of E_{ij} , as in Example 3 of Section 3.5. First, we split d_{ij} and into d_1 and d_2 which are formed by restricting d_{ij} to each of the sub-elements: i.e., d_k is d_{ij} restricted to $\gamma_{ij}^k = \gamma_{ij} \cap T_{ij}^k, k \in \{1, 2\}$;

$$d_{ij}(\eta_p) = \sum_{k \in \{1, 2\}} d_k(\eta_p), \quad (3.31)$$

$$d_k(\eta_p) \equiv - \int_{\gamma_{ij}^k} D\nabla \eta_p(u) \cdot \mathbf{n}_{ij} dS. \quad (3.32)$$

Second, we affine map from each sub-element T_{ij}^k to a reference or “work” element \widehat{T}_{ij}^k and transform the bounded linear functionals d_1 and d_2 : recall

that we lose a power of ρ when we transform first-order derivatives of u ;

$$|d_k(u)| \leq C \rho_k^{-1} \frac{|\gamma_{ij}^k|}{|\widehat{\gamma}_{ij}^k|} |\widehat{d}_k(\widehat{u})|, \quad k \in \{1, 2\}, \quad (3.33)$$

where ρ_k is the diameter of the circle inscribed in T_{ij}^k . Third, we bound $\widehat{d}_k(\widehat{u})$ and find that the Bramble-Hilbert lemma holds when the regularity parameter S for \widehat{u} satisfies $3/2 < S + 1 \leq p + 1$: the lower and upper bounds are due to the Sobolev trace theorem and to the order of the interpolation (i.e., p -th order interpolation is exact for polynomials of degree $\leq p$), respectively;

$$|\widehat{d}_k(\widehat{\eta}_p)| \leq C |\widehat{u}|_{S+1, T_{ij}^k}, \quad S \in (1/2, p]. \quad (3.34)$$

Fourth, we transform back to each sub-element and find estimates for d_1 and d_2 for $S \in (1/2, p]$:

$$\begin{aligned} |d_k(\eta_p)| &\leq C \frac{h_k^{S+1}}{\rho_k} \frac{|\gamma_{ij}^k|}{|T_{ij}^k|} |u|_{S+1, T_{ij}^k}, \\ &\leq C (h/p)^S \left(\frac{|\gamma_{ij}^k|}{|X_{ij}|} \right)^{1/2} |u|_{S+1, T_{ij}^k}, \end{aligned} \quad (3.35)$$

where we have used the facts that $h_k = O(h/p)$ is the diameter of $T_p^k \supset T_{ij}^k$, $|T_{ij}^k| \geq C |\gamma_{ij}^k| |X_{ij}|$ for some C , and the shape regularity of $\widetilde{\mathcal{T}}^{h/p}$ (i.e. $h_k/\rho_k \leq \sigma$). Note that (3.35) is an $H^1(\Omega)$ -equivalent optimal-order local estimate on the auxiliary triangulation $\widetilde{\mathcal{T}}^{h/p}$. Since $\widetilde{\mathcal{T}}^{h/p} \subseteq \mathcal{T}^{h/p} \subseteq \mathcal{T}^h$ (cf. (2.6) in Section 2.2.3)—the triangulations are nested or refinements of one another, this estimate on $\widetilde{\mathcal{T}}^{h/p}$ holds on the FVE triangulation $\mathcal{T}^{h/p}$ and the FE triangulation \mathcal{T}^h , as demonstrated below.

Combining the results for d_1 and d_2 from (3.35), we find a local error estimate for the bounded linear functional (3.19) for $S \in (1/2, p]$:

$$|d_{ij}(\eta_p(u))| \leq \sum_{k \in \{1, 2\}} |d_k(\eta_p)|,$$

$$\begin{aligned}
&\leq C (h/p)^S \left(\sum_{k \in \{1,2\}} \frac{|\gamma_{ij}^k|}{|X_{ij}|} \right)^{1/2} \left(\sum_{k \in \{1,2\}} |u|_{S+1, T_{ij}^k}^2 \right)^{1/2}, \\
&\leq C (h/p)^S \left(\frac{|\gamma_{ij}|}{|X_{ij}|} \right)^{1/2} \left(\sum_{k \in \{1,2\}} |u|_{S+1, T_{ij}^k}^2 \right)^{1/2}, \tag{3.36}
\end{aligned}$$

where we have used the fact that γ_{ij}^1 and γ_{ij}^2 partition γ_{ij} .

To obtain the final result (3.30), substitute the local estimate (3.36) into the diffusion functional (3.18) and simplify: focusing on the H^{S+1} semi-norm of u , we have

$$\begin{aligned}
\sum_{\{i,j\} \in W} \sum_{k \in \{1,2\}} |u|_{S+1, T_{ij}^k}^2 &= \sum_{T_p \in \mathcal{T}^{h/p}} |u|_{S+1, T_p}^2 = \sum_{T \in \mathcal{T}^h} |u|_{S+1, T}^2, \\
&\equiv |u|_{S+1, \mathcal{T}^h}^2 \leq \|u\|_{S+1, \Omega}^2, \tag{3.37}
\end{aligned}$$

where we've used the nested property of the triangulations (2.6): each element of $\widetilde{\mathcal{T}^{h/p}}$ is contained in a single element of $\mathcal{T}^{h/p}$ which in turn is contained in a single element of \mathcal{T}^h . \square

Following Remark 17.2 of Ciarlet [11], we offer an important corollary to Theorem 3.1.

Corollary 3.1 If $u \in C^0(\overline{\Omega}) \cap H^{S+1}(\mathcal{T}^h) \cap H^1(\Omega)$ for $S \in (1/2, p]$, i.e., u is only piecewise $S + 1$ -regular on the triangulation \mathcal{T}^h , we have the following estimate:

$$\mathcal{D}(\eta_p(u)) \leq C (h/p)^S |u|_{S+1, \mathcal{T}^h}, \quad S \in (1/2, p], \tag{3.38}$$

where

$$|u|_{S+1, \mathcal{T}^h} \equiv \left(\sum_{T \in \mathcal{T}^h} |u|_{S+1, T}^2 \right)^{1/2}. \tag{3.39}$$

Proof: Follow the proof of Theorem 3.1 up to the penultimate line of (3.37). We see that $u \in C^0(\overline{\Omega}) \cap H^{S+1}(\mathcal{T}^h) \cap H^1(\Omega)$ is the necessary for the proof to hold: $C^0(\overline{\Omega})$ is the necessary so that Lagrange interpolation is well-defined; $H^{S+1}(\mathcal{T}^h) \cap H^1(\Omega)$ is necessary to ensure that $\mathcal{D}(\cdot)$ and the piecewise semi-norm $|\cdot|_{S+1, \mathcal{T}^h}$ are bounded. \square

The result (3.38) of the corollary avoids evaluating high-order derivatives of u across the edge X_{ij} which may coincide with ∂T for an element T in \mathcal{T}^h . That is, if u is only piecewise $S + 1$ -regular on \mathcal{T}^h , $\partial^\alpha u$ may have jump discontinuities on ∂T for $|\alpha| = 1$ that generate Dirac delta functions on ∂T for $|\alpha| \geq 2$.

We work through an important illustrative application of Corollary 3.1 that will be used to derive a discrete ellipticity result in Section 3.7.

Lemma 3.4 Let $u \in \mathcal{S}^p(\mathcal{T}^h)$ and $p \geq 2$. Then the interpolation error into $\mathcal{S}^1(\mathcal{T}^{h/p})$ satisfies the following estimate:

$$\mathcal{D}(u - \Pi_1^{h/p} u) \leq C (h/p) |u|_{2, \mathcal{T}^h}. \quad (3.40)$$

Proof: For $p \geq 2$, $u \in \mathcal{S}^p(\mathcal{T}^h) \subset C^0(\overline{\Omega}) \cap H^2(\mathcal{T}^h) \cap H^1(\Omega)$ and the proof of Corollary 3.1 implies the following estimate for linear interpolation on \mathcal{T}^h :

$$\mathcal{D}(u - \Pi_1^h u) \leq C h |u|_{2, \mathcal{T}^h}.$$

Since $\mathcal{T}^{h/p}$ is a p -fold refinement of \mathcal{T}^h , the same proof holds for linear interpolation on $\mathcal{T}^{h/p}$ with h replaced by h/p : i.e.,

$$\mathcal{D}(u - \Pi_1^{h/p} u) \leq C (h/p) |u|_{2, \mathcal{T}^{h/p}},$$

where

$$|u|_{2, \mathcal{T}^{h/p}} \equiv \left(\sum_{T_p \in \mathcal{T}^{h/p}} |u|_{2, T_p}^2 \right)^{1/2}.$$

The result (3.40) follows after noting that $|u|_{2, T}^2 = \sum_{T_p \in T} |u|_{2, T_p}^2$. \square

3.6.2 Superconvergent Approximation

For linear interpolation in two spatial dimensions, we have a choice (due to Cai *et al.* [9, 10] and Süli [50]) between an $H^1(\Omega)$ -equivalent optimal-order result and a superconvergent result that depends on whether the volumization $\mathcal{V}^{h/p}$ is symmetric to the triangulation $\mathcal{T}^{h/p}$ (see Section 2.2.2). That is, from the perspective of the FVE diffusion functional, the approximating power of a linear trial space on a symmetric volumization matches the power of a quadratic trial space on a regular volumization. At the core of the superconvergence theory for the diffusive flux functional is the 1-d fact that centered differencing is exact for quadratics; we invoke certain volume symmetries, discussed below, to extend this 1-d result to multi-dimensional settings.

Theorem 3.2 (Cai and Süli) Let $\eta_1(u) = \Pi_1^h u - u$ be the linear interpolation error for $u \in \mathcal{W}^+$ into $\mathcal{S}_0^1(\mathcal{T}^h)$. Then we have the following estimate:

$$\mathcal{D}(\eta_1(u)) \leq C h^S \|u\|_{S+1, \Omega}, \quad S \in (1/2, P], \quad (3.41)$$

where $P = 1$ (optimal-order) or 2 (superconvergent) in the absence or presence of volume symmetry, respectively.

Proof: A detailed proof is given in Cai [10] that follows the steps outlined in the proof of Theorem 3.1 for $p = 1$; instead, we choose to focus on

step three where we determine the extent (i.e., the range of regularities for u) to which the Bramble-Hilbert lemma can be used to bound the diffusion functional \widehat{d}_{ij} of (3.19). We work on the whole of the “macro-element” \widehat{E}_{ij} by piecing together partial results from each \widehat{T}_{ij}^k in \widehat{E}_{ij} to obtain the superconvergence result (3.41) for symmetric volumes. That is, if \widehat{u} is quadratic on $\widehat{T}_p^1 \cup \widehat{T}_p^2$ (recall that T_p^k is the element of $\mathcal{T}^{h/p}$ that contains the element T_{ij}^k of $\widetilde{\mathcal{T}}^{h/p}$) and then restricted to \widehat{E}_{ij} , the sum of the linear interpolation errors from \widehat{T}_{ij}^1 and \widehat{T}_{ij}^2 (recall that interpolation is defined on T_p^k and then restricted to T_{ij}^k in our analysis) for the normal derivative of u will vanish if the volume is X and γ symmetric (see Section 2.2.2 for the definition of these symmetries).

The essential ideas behind this result can be seen in the following argument that holds when the diffusion coefficient is constant on γ_{ij} . Dropping the $\widehat{\cdot}$ notation for convenience, we work on a reference macro-element E_{ij} arising from a symmetric volumization (see Figure 3.1): specifically, let $X_{ij} = [-1, 1] \times 0$ and $\gamma_{ij} = 0 \times [-1, 1]$ so that X_{ij} and γ_{ij} are perpendicular bisectors of one another; in this case, the normal derivative of u on γ_{ij} is simply $\partial_x u$.

If u is quadratic, $\partial_x u$ is in the span of $\{1, x, y\}$ on E_{ij} and $\partial_x \Pi_1^h u$ is in the span of $\{1\}$ on each T_{ij}^k . Now we work on the common edge X_{ij} of the T_{ij}^k sub-elements, Using the X -symmetry of the volume, $\partial_x \Pi_1 u$ is the standard one-dimensional central difference approximation to $\partial_x u$ at $(0, 0)$.

Since central differencing is exact for quadratics by a reflection symmetry, we have the following result:

$$\partial_x \eta_1(u)|_{(0,0)} = 0. \tag{3.42}$$

Figure 3.1: Formation of the macro-element $E_{ij} = T_{ij}^1 \cup T_{ij}^2$.

In other words, evaluation of the derivative at the midpoint of an interval suppresses the midpoint DOF (the midpoint quadratic basis function has zero slope at the midpoint). Combining (3.42) with the continuity of the interpolation on X_{ij} , we find that

$$\partial_x \eta_{\mathbf{1}}(u)|_{\gamma_{ij}} \in \text{span}\{y\}. \quad (3.43)$$

Therefore, if the volume is also γ -symmetric, the functional $d_{ij}(\eta_{\mathbf{1}})$ will vanish. That is, the integrand of (3.19) will be an odd function on a symmetric γ_{ij} interval. Alternately, the midpoint or centroid quadrature rule applied to (3.19) (which gives 0 as a result) will be exact in this case when the integrand is linear. Therefore, the Bramble-Hilbert lemma holds when the regularity parameter S is in $(1/2, 2]$ and the result (3.41) holds.

Notice that if bilinear interpolation on a non-uniform tensor-product mesh is employed (i.e., we use a bilinear trial space to represent the FVE solution), as in Süli [50], only X -symmetry is required for superconvergence. That is, if u is quadratic, $\partial_x \Pi_{1,1}^h u$ is in the span of $\{1, y\}$. Using the X -symmetry of the volume, we see that the one-dimensional central differencing or reflection symmetry result of (3.42) extends to any $y \in \gamma_{ij}$ under bilinear interpolation:

$$\partial_x \eta_{1,1}(u)|_{(0,y)} = 0; \quad (3.44)$$

i.e., for any fixed value of y the previous result (3.42) holds. Therefore, the integrand of (3.19) vanishes on any γ_{ij} interval. The key to success for bilinears is that they contain the quadratic xy term that is sufficient to approximate the normal derivative exactly on γ_{ij} when X -symmetry is present.

To demonstrate superconvergence for a variable diffusion coefficient, we use a piecewise constant approximation $\bar{D} \equiv \Pi_0^h D$ on γ_{ij} . Then with the modified notation $d_{ij}(u) \equiv d_{ij}(u; D)$, we consider the splitting induced by \bar{D} :

$$d_{ij}(\eta_1(u); D) = d_{ij}(\eta_1(u); \bar{D}) + d_{ij}(\eta_1(u); \eta_0(D)). \quad (3.45)$$

The first term on the right-hand side of (3.45) is estimated by the linear Bramble-Hilbert lemma as indicated above. However, the second term on the right-hand side of (3.45) requires the bilinear Bramble-Hilbert lemma of Ciarlet [11]. The bilinear lemma allows us to obtain an $O(h^2)$ estimate by multiplying the $O(h)$ estimates (determined by the linear lemma) for linear interpolation of $\partial_x u$ and constant interpolation of D . \square

3.6.2.1 Extension to Three Dimensions

The proof of Theorem 3.2 is based on volume symmetries in two spatial dimensions ($d = 2$); however, we will see that the notion of volume symmetry is dimensionally and geometry dependent. If $d = 1$, the modifications in the proof are trivial: γ_{ij} is a point in 1-d and $|\gamma_{ij}|$ is replaced by 1 as stated in Section 3.3, so the only volume symmetry is X -symmetry: i.e., the volume boundary γ_{ij} is at the midpoint of each element edge X_{ij} ; in this simple case, the diffusive flux is represented by a centered difference quotient.

In three spatial dimensions, the modifications in the proof are non-trivial: below we outline necessary steps to extend the result (3.41) to this important case.

First, we upgrade our definition of volume symmetry: for X -symmetry,

we require the planar surface γ be an orthogonal bisector for X_{ij} ; for “centroid” γ -symmetry, we require $X_{ij} \in \gamma_{ij}$ and for the intersection point P_{ij} to be the centroid or center of mass for γ_{ij} ; furthermore, we require that the geometry of γ_{ij} is such that the centroid quadrature rule is exact for linear functions: e.g., γ_{ij} is rectangular, triangular, hexagonal, etc. (see Stroud [49] for details).

Second, we consider the use of a C^0 linear trial space on tetrahedra. If u is quadratic, $\partial_x u$ is in the span of $\{1, x, y, z\}$ on E_{ij} and $\partial_x \Pi_1^h u$ is in the span of $\{1\}$ on each T_{ij}^k . Using the X -symmetry of the volume, the one-dimensional central differencing or reflection symmetry (3.42) result becomes:

$$\partial_x \eta_1(u)|_{(0,0,0)} = 0. \quad (3.46)$$

Combining (3.46) with the continuity of the interpolation on element faces that have X_{ij} as an edge, we find that

$$\partial_x \eta_1(u)|_{\gamma_{ij}} \in \text{span}\{y, z\}. \quad (3.47)$$

Therefore, if the volume is also γ -symmetric, the functional $d_{ij}(\eta_1)$ will vanish. That is, the centroid quadrature rule (which gives 0 as result) will be exact. Therefore, the Bramble-Hilbert lemma holds when the regularity parameter S is in $(1/2, 2]$ and the result (3.41) holds in three dimensions under assumptions of X and γ volume symmetries, as it did in two dimensions.

However, for C^0 bilinears (i.e., $\Pi_{1,1} u \in \text{span}\{1, x, y\} \otimes \{1, z\}$) on triangular prisms formed in a tensor-product mesh, volume symmetry requires X -symmetry and a “2-d” γ -symmetry: i.e., the x and y cross-sections of γ_{ij} must possess the centroid symmetry. That is, linearity in x and y requires

the 2-d γ_{ij} -symmetry in x and y , but bilinearity in z eliminates the 2-d γ_{ij} -symmetry in z . The analog of (3.44) is

$$\partial_x \eta_{1,1}(u)|_{(0,0,z)} = 0; \quad (3.48)$$

the bilinearity in z eliminates γ symmetries in the z -direction as in the 2-d bilinear case. However, the residue from the linearity in y remains:

$$\partial_x \eta_{1,1}(u)|_{\gamma_{ij}} \in \text{span}\{y\}; \quad (3.49)$$

in order to eliminate this term, we require that the y cross-section of γ_{ij} possess the centroid symmetry as in the 2-d linear case. The same remarks hold for the analysis of $\partial_y \eta_{1,1}$; the analysis of $\partial_z \eta_{1,1}$ follows the 2-d bilinear case exactly.

Finally, for C^0 trilinears on rectangular bricks formed by a tensor-product mesh, volumesymmetry requires X -symmetry and a “1-d” γ_{ij} -symmetry: i.e., either one of the 1-d cross-sections of γ_{ij} must possess the centroid symmetry. That is, the quadratic components of a trilinear trial space obviate γ_{ij} -symmetry requirements as in 2-d; however, its cubic component invokes a reduced, 1-d γ_{ij} -symmetry. A generalization of (3.44),

$$\partial_x \eta_{1,1,1}(u)|_{(0,y,0),(0,0,z)} = 0, \quad (3.50)$$

eliminates the quadratic components of the trilinear interpolant. However, the residue from the cubic component remains:

$$\partial_x \eta_{1,1,1}(u)|_{\gamma_{ij}} \in \text{span}\{yz\}. \quad (3.51)$$

In order to eliminate this term, we require that either a y or z cross-section of γ_{ij} possess the centroid symmetry. So trilinear superconvergence requires a “reduced” γ -symmetry in addition to X -symmetry.

Unlike standard FE superconvergence results that are usually local and one-dimensional, the FVE superconvergence result (3.41) is always global and multidimensional. Notice (cf. (3.30)) that the superconvergence is consistent with the use of a quadratic, rather than a linear, trial space.

3.6.2.2 Degenerate Volume Paradigm

To help understand the robustness of the superconvergence result (3.41) in two and three dimensions, we note the the linear FVE on a symmetric volumization is identical to a quadratic FVE (using a one-point centroid quadrature rule for the diffusion functional (3.19) at the intersection of X_{ij} and γ_{ij}) on a “degenerate” volumization posed on a uniform tensor-product mesh. That is, if we allow the volumes about the midpoint DOF in \mathcal{T}^h to degenerate to sets of measure zero (see Figure 3.2), only vertex DOF in \mathcal{T}^h are expressed in the resulting matrix system—all of the midpoint DOF are systematically suppressed as can be seen from the one-dimensional reflection symmetry results (3.42) and (3.46). Also, the error in the one-point centroid quadrature is $O(h^2)$ and will not impair the superconvergence, as can be seen in the analysis of the effects of numerical quadrature in [10].

In this process, the volumization $\mathcal{V}^{h/2}$, which is consistent with the use a quadratic trial space, degenerates to the symmetric volumization \mathcal{V}^h , which is consistent with the use of a linear trial space. As can be easily verified by direct calculation using the (3.43) and (3.47), the two methodologies yield identical matrix systems. Furthermore, if a bilinear trial space is used, the “degenerate” volume paradigm holds on a non-uniform tensor-product mesh (without the use of a centroid quadrature rule) as can be seen from (3.44).

Figure 3.2: An example of $\mathcal{V}^{h/2}$ degenerating to \mathcal{V}^h .

Therefore, the symmetric volume method for linear FVE can be thought of as a degenerate volume method for quadratic FVE. From this perspective, the robust superconvergence behavior for linear FVE can be understood as a special case of the “ordinary” optimal-order behavior for quadratic FVE.

Finally, it is important to emphasize the local character of the error estimates of this chapter, since it may appear that superconvergence (for linear trial spaces) is strictly relegated to uniform computational meshes and that only linear convergence is to be expected on non-uniform meshes. Recall that superconvergence results hold on the “macro-element” E_{ij} formed by joining individual element edges X_{ij} with individual volume interfaces γ_{ij} . Even when global symmetry conditions are violated, local (second-order) superconvergence holds when local symmetry conditions for X_{ij} and γ_{ij} are satisfied. Qualitatively, global superconvergence can be seen as the cumulative effect of local superconvergence behavior. Quantitatively, local superconvergence leads to superlinear global convergence rates: the precise rate depends upon the prevalence of local symmetry or local uniformity on the global computational mesh.

Therefore, on general non-uniform meshes, “partial” superconvergence can exist globally even when full second-order superconvergence is lost (see [9] for details). Even on highly non-uniform meshes, there are usually significant regions of uniformity and symmetry. In these local regions, the solution converges quadratically and “remediates” the global solution: superlinear, rather than linear, convergence is observed globally.

3.7 Ellipticity

In Section 2.6.4, use of the ellipticity result (2.23) in the FE error analysis required that the trial space representation error be used as a test function. In FE, test and trial spaces coincide so this was straightforward. In FVE, test and trial spaces differ, and we must describe the test function representation of a trial function.

Let $u \in \mathcal{W}^h$ with nodal representation $u \equiv \sum_{i \in I} u_i \phi_i$, where $\{\phi_i\}_{i \in I}$ is the DOF basis for \mathcal{W}^h . Then u also has a corresponding *lumped* or test space representation, \bar{u} , in \mathcal{X}^h :

$$\bar{u} \equiv \sum_{i \in I} u_i \chi_i. \quad (3.52)$$

In other words, if $u \in \mathcal{S}^p(\mathcal{T}^h)$, $\bar{u} \in \mathcal{S}^0(\mathcal{V}^{h/p})$ is a volume-wise constant interpolant of u : $\bar{u} = \Pi_0^{h/p} u$. Our ellipticity result for FVE will use the lumped representation defined by (3.52) to characterize test functions.

As in FE, a discrete ellipticity result is perhaps the most important technical component of FVE analysis. Once ellipticity is established, FVE error analysis for diffusion equation reduces to the approximation theory developed in the previous section.

Using the optimal-order approximation result (3.40) of Lemma 3.4 from Section 3.6, we derive (for $p \geq 2$) an “asymptotic” ellipticity result for the diffusion bilinear form in terms of our discrete FVE trial space $\mathcal{S}^p(\mathcal{T}^h)$ and our discrete H^1 semi-norm. That is, for (h/p) sufficiently small, the ellipticity constant β is independent of the computational mesh. This type of argument is unavoidable if we wish to maintain the FE formalism in our FVE error analysis

for higher-order ($p \geq 2$) polynomial trial spaces, since our discrete H^1 semi-norm is defined as the restriction of the continuous H^1 semi-norm to piecewise linear polynomials on the FVE triangulation $\mathcal{T}^{h/p}$.

Lemma 3.5 (Discrete Ellipticity) There exists a positive constant $\beta = \beta(D_m; \lambda_D)$, for $h/p \leq \lambda_D$ when $p \geq 2$, such that

$$B(u, \bar{u}) \geq \beta |u|_{1,\omega}^2, \quad \forall u \in \mathcal{W}^h, \quad (3.53)$$

where D_m is the positive lower bound on the diffusion coefficient and $\lambda_D = \frac{1}{2} D_m C^{-1} |u|_{1,\omega} / |u|_{2,\mathcal{T}^h}$ is an asymptotic ellipticity parameter.

Proof: For $u \in \mathcal{S}_0^1(\mathcal{T}^{h/p})$, we follow Cai[10] for the following sub-result. By the definition of the B -form (3.8), we have

$$\begin{aligned} B(u, \bar{u}) &= \sum_{i \in I} u_i \sum_{j \in W_i} d_{ij}(u), \\ &= \sum_{\{i,j\} \in W} (u_i - u_j) d_{ij}(u), \end{aligned}$$

since $d_{ij}(u) = -d_{ji}(u)$ (i.e., $\mathbf{n}_{ij} = -\mathbf{n}_{ji}$).

For any $u \in \mathcal{S}_0^1(\mathcal{T}^{h/p})$, we recognize that

$$d_{ij}(u) = \frac{(u_i - u_j)}{|X_{ij}|} \int_{\gamma_{ij}} D dS.$$

Then, it follows that

$$\begin{aligned} B(u, \bar{u}) &= \sum_{\{i,j\} \in W} \frac{(u_i - u_j)^2}{|X_{ij}|} \int_{\gamma_{ij}} D dS, \\ &\geq D_m \sum_{\{i,j\} \in W} (u_i - u_j)^2 \frac{|\gamma_{ij}|}{|X_{ij}|}, \\ &= D_m |u|_{1,\omega}^2. \end{aligned} \quad (3.54)$$

Thus, the lemma is proved for the $p = 1$ case with $\beta = D_m$.

For $u \in \mathcal{S}^p(\mathcal{T}^h)$ with $p \geq 2$, define $u_1 \equiv \Pi_1^{h/p} u$. Now, analyze the following splitting induced by u_1 :

$$\begin{aligned}
B(u, \bar{u}) &= B(u_1, \bar{u}_1) + B(u - u_1, \bar{u}), \\
&\geq D_m |u|_{1,\omega}^2 - \mathcal{D}(u - u_1) |u|_{1,\omega}, \\
&\geq D_m |u|_{1,\omega}^2 - C(h/p) |u|_{2,\mathcal{T}^h} |u|_{1,\omega}, \\
&\geq \frac{1}{2} D_m |u|_{1,\omega}^2,
\end{aligned} \tag{3.55}$$

where $h/p \leq \lambda_D = \frac{1}{2} D_m C^{-1} |u|_{1,\omega} / |u|_{2,\mathcal{T}^h}$.

In the first three lines of (3.55), we use the fact that $\bar{u} = \bar{u}_1$, the sub-result (3.54), the fact that $|u_1|_{1,\omega} = |u|_{1,\omega}$, the bound (3.17), and the approximation result (3.40), respectively. Thus, the lemma is proved for the $p \geq 2$ case with $\beta = \frac{1}{2} D_m$. \square

In addition, we see that $B(u, \bar{u}) \equiv \mathbf{u}^T \mathbf{B} \mathbf{u} \equiv Q_{\mathbf{B}}(\mathbf{u}) \quad \forall u \in \mathcal{W}^h$, where \mathbf{u} is the vector of nodal values for $u \in \mathcal{W}^h$, \mathbf{B} is the FVE matrix generated by the B -form, and $Q_{\mathbf{B}}(\cdot)$ is the associated quadratic form. Therefore, (3.53) simply states that \mathbf{B} is positive definite:

$$Q_{\mathbf{B}}(\mathbf{u}) \geq 0, \quad \forall \mathbf{u}, \tag{3.56}$$

with equality holding in (3.56) only when \mathbf{u} (i.e., $u \in \mathcal{W}^h$) is identically zero. Hence, (3.53) guarantees the existence and uniqueness of the numerical solution $u^h \in \mathcal{W}^h$.

The FVE ellipticity result (3.53) clearly has the same form as the corresponding FE result (2.23) (when restricted to \mathcal{W}^h). In fact, the two

results are identical for (2.14) when D is constant and $\mathcal{W}^h = \mathcal{S}_0^1(\mathcal{T}^h)$; that is, the FVE (and FV) matrix \mathbf{B} and the FE matrix \mathbf{A} are identical for piecewise linear polynomials on a general triangulation as demonstrated by Bank and Rose [5]. This equivalence between FV and FE for a piecewise linear trial space is the key observation that drives much of the FV analysis in the references cited in Section 1.1. Such arguments may seem to imply dependence of FV (or FVE) analysis on the corresponding FE formulation of the same problem. As we have seen in the preceding sections (and will see again in later sections), this is not the case: our FVE analysis is similar to FE in style *but not* in substance.

3.8 Error Analysis

With the components developed in the preceding sections, expressed in (3.17), (3.30), (3.41), and (3.53), we can outline the standard FVE analysis. To combine optimal-order and superconvergence results, define

$$\rho \equiv \max\{p, P\}, \quad (3.57)$$

where p and P are the optimal-order and superconvergence parameters employed in Theorem 3.2. The pathway to our error estimate is detailed below.

3.8.1 Error Splitting

As in the FE analysis of Section 2.6.4, we define the error in the FVE approximation to (2.14) as $e \equiv u - u^h \in \mathcal{W}$ and write $e = \zeta - \eta_p(u)$, so that the FVE error is split into its representation error ($\zeta \equiv \Pi_p^h u - u^h \in \mathcal{W}^h$) and interpolation error ($\eta_p(u) \equiv \Pi_p^h u - u \in \mathcal{W}$) components. Since the interpolation error is zero at all computational mesh points ω , $|\eta_p(u)|_{1,\omega} \equiv 0$ and $|u - u^h|_{1,\omega} \equiv |\zeta|_{1,\omega}$. Hence, the $H^1(\omega)$ semi-norm focuses our attention on the representation error of the discrete trial space \mathcal{W}^h for our numerical method and on the behavior of the error on the points ω where the numerical approximation is being computed rather than on the whole domain Ω or on the continuous trial space \mathcal{W}^+ .

Therefore, estimating $|e|_{1,\omega}$ is reduced to estimating $|\zeta|_{1,\omega}$; however, the latter estimate can be developed as below in terms of estimates (3.30) and (3.41) for the interpolation error η_p over the set of volume interfaces $\{\gamma_{ij}\}_{ij \in \mathcal{W}}$ that are contained in $\overline{\Omega}$.

3.8.2 Zero Property

Knowing that (for all $w \in \mathcal{X}^h$) the exact solution of (3.1) satisfies $B(u, w) = (f, w)$, the numerical solution satisfies (3.7), and the source term satisfies (3.10), we have the *zero property*,

$$B(u - u^h, w) = 0, \quad \forall w \in \mathcal{X}^h, \quad (3.58)$$

which in our context is more appropriately expressed as

$$B(\zeta, w) = B(\eta_p, w), \quad \forall w \in \mathcal{X}^h. \quad (3.59)$$

3.8.3 Error Estimate

Since $\bar{\zeta}$, the lumped representation (or $\mathcal{S}^0(\mathcal{V}^{h/p})$ -interpolant) of ζ , is in \mathcal{X}^h , take $w = \bar{\zeta}$ in (3.59) and obtain

$$\begin{aligned} \beta |\zeta|_{1,\omega}^2 &\leq B(\zeta, \bar{\zeta}), \\ &= B(\eta_p, \bar{\zeta}), \\ &\leq \mathcal{D}(\eta_p) |\zeta|_{1,\omega}. \end{aligned} \quad (3.60)$$

The three lines of (3.60) use (3.53), (3.59), and (3.17), respectively. Finally, by (3.30) and (3.41), we have proved the following combined optimal-order and superconvergent error estimate for the finite volume element method. \square

Theorem 3.3 (Error Estimate) The error in the FVE approximation (3.4) to the solution of the diffusion equation (2.14) satisfies

$$|u - u^h|_{1,\omega} \leq C (h/p)^S \|u\|_{S+1,\Omega}, \quad S \in (1/2, \rho], \quad (3.61)$$

where ρ is defined in (3.57).

Note that (3.61) matches or surpasses the corresponding p th-order finite element result (2.27).

This concludes our presentation for diffusion equations. In the next chapter, we outline the necessary additions and modifications for the inclusion of reaction and advection terms, as in the general elliptic equation (1.1) that arises from the general integral conservation law (1.4) of Section 1.2.

3.9 Summary

In the present chapter, we have followed the formalisms of the h , p , and h - p version FE error analysis presented in Chapter 2 to develop an h , p , and h - p version FVE analysis for diffusion equations. In Sections 3.1 and 3.2, we motivated and laid the groundwork for an FE-style analysis of FVE. In Section 3.3, we developed discrete analogs of the L^2 and H^1 Sobolev norms and a discrete Poincaré inequality as the first step in maintaining the Sobolev formalism of Section 2.5 in our FVE analysis. In Section 3.4, we derived an upper bound for the FVE diffusion bilinear forms in terms of our discrete H^1 semi-norm and an FVE diffusion functional by following in the footsteps of Section 2.6.1. In Section 3.5, we introduced the Bramble-Hilbert lemma as a new tool for the development of an FVE approximation theory in Section 3.6.

Following the footsteps of Section 2.6.2, $H^1(\Omega)$ -equivalent optimal-order error estimates were derived in the $H^1(\omega)$ semi-norm for h , p , and h - p version FVE in Section 3.6.1. In Section 3.6.2, the proofs of the two dimensional superconvergence results of Cai *et al.* [9, 10] for a linear trial space on triangles and Süli [50] for a bilinear trial space on rectangles were outlined: on a “symmetric” volumization, the linear FVE has the same asymptotic convergence rate as quadratic FVE on a regular volumization. We found that superconvergence for the diffusive flux is based the 1-d fact that centered differencing is exact for quadratics: various volume symmetries are introduced to extend this result to multi-dimensional settings. This set the stage for Section 3.6.2.1 where the dimensional and geometric dependence of volume symmetry and superconvergence was unmasked in three spatial dimensions. Finally, we

offered the “degenerate” volume paradigm” in Section 3.6.2.2 to show that linear FVE on a symmetric volumization is identical to a quadratic FVE on a degenerate volumization; hence, superconvergence for linear FVE can be viewed as a special case of the optimal-order convergence for quadratic FVE.

When discussing superconvergence for linear trial spaces, it is important to emphasize that it is primarily a local phenomenon. Qualitatively, global superconvergence is the cumulative effect of local volume symmetries or uniformities in a given computational mesh. Quantitatively, local quadratic convergence leads to superlinear global convergence: the precise rate depends upon the prevalence of local volume symmetry or uniformity conditions on a global computational mesh. As a result, local quadratic convergence of a numerical solution (in regions of symmetry/uniformity on a non-uniform mesh) “remediates” the global solution: superlinear, rather than linear, convergence is observed globally.

Following the formalism of Section 2.6.3, we used the optimal-order approximation result (3.40) of Lemma 3.4 from Section 3.6.1 to derive (for $p \geq 2$) an “asymptotic” ellipticity result for the diffusion bilinear form in terms of our discrete FVE trial space $\mathcal{S}^p(\mathcal{T}^h)$ and our discrete H^1 semi-norm in Section 3.7. That is, for (h/p) sufficiently small, the ellipticity constant is independent of the computational mesh. This type of argument is unavoidable if we wish to maintain the FE formalism in our FVE error analysis for higher-order ($p \geq 2$) polynomial trial spaces, since our discrete H^1 semi-norm is defined as the restriction of the continuous H^1 semi-norm to piecewise linear polynomials on the FVE triangulation $\mathcal{T}^{h/p}$. Following the pattern of Section

2.6, we performed an FE-style error analysis for h , p , and h - p version FVE in Section 3.8 that matches or exceeds the results of the standard h , p , and h - p FE analysis for steady diffusion equations.

In the next chapter, we extend the scope of our analysis to general elliptic equations and detail the analogous theoretical results for the FVE treatment of advection and reaction.

4. Analysis for General Elliptic Equations

In this chapter, we note modifications for an h , p , and h - p version FVE analysis for general elliptic equations and elliptic projections. After presenting some preliminary theoretical infrastructure in Section 4.1, we develop upper bounds in terms of FVE advection and reaction functionals and our discrete L^2 and H^1 norms for the FVE bilinear forms corresponding to the treatment of advection and reaction terms in a general elliptic equation in Section 4.2. In Section 4.3, we develop L^2 -equivalent optimal-order approximation theory results for our FVE advection and reaction functionals. Our goal here is to find estimates for advection and reaction that do not interfere or add any additional “volume symmetry” constraints to the diffusion optimal-order and superconvergence results of the previous chapter. In Section 4.4, we derive “asymptotic” discrete ellipticity results for the FVE reaction and advection bilinear forms. Finally in Section 4.5, we construct an FVE analysis for reaction–diffusion and advection–reaction–diffusion equations with the components developed in the preceding sections: optimal-order and superconvergence results are obtained for h , p , and h - p version FVE in our discrete H^1 semi-norm. In Section 4.6, we extend the theory developed so far to analyze the elliptic map which is the key theoretical tool in our analysis of parabolic equations.

4.1 Preliminaries

In general, the FVE B -form is given by

$$B(u, w) \equiv \mathcal{A}(u, w) + \mathcal{D}(u, w) + \mathcal{R}(u, w), \quad u \in \mathcal{H}, w \in \mathcal{X}^h, \quad (4.1)$$

where the bilinear forms corresponding to advection, diffusion, and reaction are

$$\mathcal{A}(u, w) \equiv \sum_{i \in I} \int_{\partial V_i} \mathbf{a} u \cdot \mathbf{n} w_i dS, \quad (4.2)$$

$$\mathcal{D}(u, w) \equiv \sum_{i \in I} - \int_{\partial V_i} D \nabla u \cdot \mathbf{n} w_i dS, \quad (4.3)$$

$$\mathcal{R}(u, w) \equiv \sum_{i \in I} \int_{V_i} r u w_i d\mathbf{x}. \quad (4.4)$$

Recall that (4.2)–(4.4) correspond to the left-hand side of the general integral equation posed (1.4) on the volumization $\mathcal{V}^{h/p}$, so that (4.1) is the starting point for an FVE analysis for the general elliptic equation (1.1).

Since the exact solution u satisfies (1.4) on a arbitrary set of volumes, it also satisfies (1.4) on the finite set of volumes $\mathcal{V}^{h/p}$ according to (4.1): given $f \in L^2$, $u \in \mathcal{W}^+$ satisfies

$$B(u, w) = (f, w), \quad \forall w \in \mathcal{X}^h. \quad (4.5)$$

As before, the right-hand side of (4.5) is given by

$$(f, w) = \sum_{i \in I} \int_{V_i} f w_i d\mathbf{x}. \quad (4.6)$$

Now we can define the FVE approximation u^h to (1.4) in terms of (4.5) and (4.6): given $f^h \in \mathcal{W}^h$, find $u^h \in \mathcal{W}^h$ such that

$$B(u^h, w) = (f^h, w), \quad \forall w \in \mathcal{X}^h. \quad (4.7)$$

As before, the numerical source term f^h satisfies the local L^2 projection relation,

$$(f - f^h, w) = 0, \quad \forall w \in \mathcal{X}^h, \quad (4.8)$$

which ensures that (source term) mass is conserved on each volume.

In Sections 3.4–3.7, we analyzed the diffusion bilinear form (4.3); here we outline the corresponding results for the advection and reaction forms (4.2) and (4.4). For reaction problems, we assume continuous and bounded ($0 < r_m \leq r \leq r_M < \infty$) reaction coefficients. For advection problems, we assume continuously differentiable and bounded ($\|\mathbf{a}\| \leq M_{\mathbf{a}}, \|\nabla \cdot \mathbf{a}\| \leq M_{\mathbf{div}} < \infty$) velocity vectors.

4.2 Boundedness

As in Section 3.4, we will develop upper bounds for the advection and reaction bilinear forms in terms of a product of functionals applied to test and trial space functions. For elements of the FVE test space \mathcal{X}^h , we apply the discrete L^2 and H^1 Sobolev norms of Section 3.3 that are linked by the discrete Poincaré inequality (3.16). For the admissible FVE trial space, $\mathcal{H} \equiv \mathcal{W}^+ \oplus \mathcal{W}^h$, we apply some problem-dependent functionals that correspond to the FVE treatment of advection and reaction. Practically speaking, we want these reaction and advection functionals to support the optimal-order or superconvergent approximation theorems of Section 3.6, since the trial function in our analysis is the interpolation error for the exact solution.

4.2.1 Advection

For the advection bilinear form (4.2), we derive a bound based on the discrete H^1 semi-norm (3.13) that is similar to the diffusion bound (3.17).

Lemma 4.1

$$|\mathcal{A}(u, w)| \leq \mathcal{A}(u) |w|_{1, \omega}, \quad \forall u \in \mathcal{H}, w \in \mathcal{X}^h. \quad (4.9)$$

$\mathcal{A}(\cdot)$ is a bounded functional:

$$\mathcal{A}(u) \equiv \left(\sum_{\{i,j\} \in W} a_{ij}^2(u) \frac{|X_{ij}|}{|\gamma_{ij}|} \right)^{1/2}; \quad (4.10)$$

$a_{ij}(\cdot)$ is a bounded linear functional arising from the FVE treatment of advection:

$$a_{ij}(u) \equiv \int_{\gamma_{ij}} \mathbf{a}u \cdot \mathbf{n}_{ij} dS, \quad (4.11)$$

where \mathbf{n}_{ij} is the unit normal vector pointing outward from V_i into V_j .

Proof: Using the definition (4.2) and the fact that

$$\int_{\partial V_i} \mathbf{a}u \cdot \mathbf{n}_i dS \equiv \sum_{j \in W_i} a_{ij}(u), \quad (4.12)$$

we follow the proof of the diffusion bound:

$$\begin{aligned} |\mathcal{A}(u, w)| &= \left| \sum_{i \in I} w_i \sum_{j \in W_i} a_{ij}(u) \right|, \\ &= \left| \sum_{\{i,j\} \in W} a_{ij}(u) (w_i - w_j) \right|, \\ &\leq \left(\sum_{\{i,j\} \in W} a_{ij}^2(u) \frac{|X_{ij}|}{|\gamma_{ij}|} \right)^{1/2} \left(\sum_{\{i,j\} \in W} (w_i - w_j)^2 \frac{|\gamma_{ij}|}{|X_{ij}|} \right)^{1/2}, \end{aligned} \quad (4.13)$$

where we have used the fact that $a_{ij}(u) = -a_{ji}(u)$ (i.e., $\mathbf{n}_{ij} = -\mathbf{n}_{ji}$) in the second line of (4.13). \square

4.2.2 Reaction

For the reaction bilinear form (4.4), we derive a bound based on our discrete L^2 norm (3.11).

Lemma 4.2

$$|\mathcal{R}(u, w)| \leq \mathcal{R}(u) |w|_{0,\omega}, \quad \forall u \in \mathcal{H}, w \in \mathcal{X}^h. \quad (4.14)$$

$\mathcal{R}(\cdot)$ is a bounded functional:

$$\mathcal{R}(u) \equiv \left(\sum_{i \in I} \frac{1}{|V_i|} r_i^2(u) \right)^{1/2}; \quad (4.15)$$

$r_i(\cdot)$ is a bounded linear functional arising from the FVE treatment of reaction:

$$r_i(u) = \int_{V_i} ru \, d\mathbf{x}. \quad (4.16)$$

Proof: Starting with the definition (4.4), we find that

$$\begin{aligned} |\mathcal{R}(u, w)| &= \left| \sum_{i \in I} r_i(u) w_i \right|, \\ &\leq \left(\sum_{i \in I} \frac{1}{|V_i|} r_i^2(u) \right)^{1/2} \left(\sum_{i \in I} |V_i| w_i^2 \right)^{1/2}, \end{aligned} \quad (4.17)$$

after an application of the discrete Cauchy-Schwarz inequality. \square

With upper bounds for the advection and reaction bilinear forms in terms of the non-standard, problem-dependent, bounded linear functionals (4.11) and (4.16), we need approximation theorems: these results are demonstrated and discussed below.

4.3 Approximation Theory

The approximation theory developed in this section forms the core of our FVE analysis for general elliptic equations in Section 4.5 and for general parabolic equations in the next chapter. Our goal here is to find estimates for advection and reaction that do not interfere or add any additional “volume symmetry” constraints to the diffusion optimal-order and superconvergence results of Theorems 3.1 and 3.2. To achieve this goal, we develop $L^2(\Omega)$ -equivalent optimal-order results for interpolation into $\mathcal{S}^p(\mathcal{T}^h)$ when the triangulation and volumization are regular (defined in Sections 2.1.1 and 2.2.3). These results provide an additional power of h/p to the H^1 optimal-order result (3.30) and coincide with the H^1 superconvergence result (3.41) for linear finite elements on any regular volumization.

4.3.1 Advection

For the advection functional (4.10), we apply the linear Bramble-Hilbert lemma to the bounded linear functional (4.11) in the same way that we applied the lemma to the bounded linear functional (3.19) to obtain an $L^2(\Omega)$ -equivalent optimal-order result for p th-order interpolation (cf. (3.30)).

Theorem 4.1 Let $\eta_p(u) = \Pi_p^h u - u$ be the interpolation error for $u \in \mathcal{W}^+$ into $\mathcal{S}^p(\mathcal{T}^h)$. Then we have the following estimate:

$$\mathcal{A}(\eta_p(u)) \leq C (h/p)^{S+1} \|u\|_{S+1, \Omega}, \quad S \in (1/2, p], \quad (4.18)$$

Proof: We follow the proof of Theorem 3.1 line-by-line with a minor modification for advection: since there is no derivative to transform in (4.11), a power of h is not lost in the affine mapping to a reference element as occurred

in the second step of the diffusion proof. Thus, we obtain a local error estimate for the bounded linear functional (4.11) for $S \in (1/2, p]$ (cf. (3.36)):

$$|a_{ij}(\eta_p(u))| \leq C (h/p)^{S+1} \left(\frac{|\gamma_{ij}|}{|X_{ij}|} \right)^{1/2} \sum_{k \in \{1,2\}} |u|_{S+1, T_{ij}^k}^2. \quad (4.19)$$

To obtain the final result (4.18), substitute the local estimate (4.19) into the advection functional (4.10) and simplify as in (3.37). \square

4.3.2 Reaction

For the reaction functional (4.15), we apply the linear Bramble-Hilbert lemma to the bounded linear functional (4.16) in a slightly different way to obtain an $L^2(\Omega)$ -equivalent optimal-order result for p th-order interpolation (cf. (3.30)).

Theorem 4.2 Let $\eta_p(u) = \Pi_p^h u - u$ be the interpolation error for $u \in \mathcal{W}^+$ into $\mathcal{S}^p(\mathcal{T}^h)$. Then we have the following estimate:

$$\mathcal{R}(\eta_p(u)) \leq C (h/p)^{S+1} \|u\|_{S+1, \Omega}, \quad S \in (1/2, p], \quad (4.20)$$

Proof: Following the volume construction in Section 2.2.1, we work on each sub-volume $v_i(T_p)$ in V_i which in turn is contained in $\mathcal{T}_i^{h/p}$ —the collection of elements T_p in $\mathcal{T}^{h/p}$ that have \mathbf{x}_i as a vertex: recall (2.3) and note that $v_i(T_p) = V_i \cap T_p$. Then we follow the proof of Example 2 from Section 3.5 for p th-order interpolation to get estimates on each sub-volume.

First, we split (4.16) into its sub-volume components:

$$r_i(\eta_p) = \sum_{T_p \in \mathcal{T}_i^{h/p}} r_{v_i(T_p)}(\eta_p), \quad (4.21)$$

$$r_{v_i(T_p)}(\eta_p) \equiv \int_{v_i(T_p)} r \eta_p(u) d\mathbf{x}. \quad (4.22)$$

Second, recalling (Section 2.2.1 and (2.5)) that interpolation is defined on each T_p in $\mathcal{T}^{h/p}$ and restricted to each $v_i(T_p)$ in $v^{h/p}$, we affine map from each element T_p to a reference element \widehat{T}_p and transform our sub-volume functionals:

$$|r_{v_i(T_p)}(\eta_p)| = \frac{|v_i(T_p)|}{|\widehat{v}_i(\widehat{T}_p)|} |\widehat{r}_{\widehat{v}_i(\widehat{T}_p)}(\widehat{\eta}_p)|. \quad (4.23)$$

Third, we bound $\widehat{r}_{\widehat{v}_i(\widehat{T}_p)}$ and find that the Bramble-Hilbert lemma holds when the regularity parameter S for \widehat{u} satisfies $3/2 < S + 1 \leq p + 1$: the upper bound is due to the order of the interpolation (i.e., p -th order interpolation is exact for polynomials of degree $\leq p$); the lower bound is chosen to mesh with the lower bound of Theorem 3.1. That is,

$$|\widehat{r}_{\widehat{v}_i(\widehat{T}_p)}(\widehat{\eta}_p)| \leq C |\widehat{u}|_{S+1, \widehat{v}_i(\widehat{T}_p)}, \quad S \in (1/2, p]. \quad (4.24)$$

Fourth, we transform back to each element T_p in $\mathcal{T}_i^{h/p}$ and find estimates for $r_{v_i(T_p)}(\eta_p)$:

$$|r_{v_i(T_p)}(\eta_p)| \leq C (h/p)^{S+1} |v_i(T_p)|^{1/2} |u|_{S+1, v_i(T_p)}, \quad S \in (1/2, p], \quad (4.25)$$

where $O(h/p)$ is the diameter of $v_i(T_p)$. Note that (4.25) is an $L^2(\Omega)$ -equivalent optimal-order local estimate on the sub-volume $v_i(T_p)$. Since $v^{h/p} \subseteq \mathcal{T}^{h/p} \subseteq \mathcal{T}^h$ (cf. (2.5))—this estimate on $v_i(T_p)$ holds on the FVE triangulation $\mathcal{T}^{h/p}$ and the FE triangulation \mathcal{T}^h , as demonstrated below.

After combining sub-volume results of (4.25), we obtain a local error estimate for the bounded linear functional (4.16) for $S \in (1/2, p]$ (cf. (3.36)):

$$|r_i(\eta_p(u))| \leq \sum_{T_p \in \mathcal{T}_i^{h/p}} |r_{v_i(T_p)}(\eta_p)|,$$

$$\begin{aligned}
&\leq C (h/p)^{S+1} \left(\sum_{T_p \in \mathcal{T}_i^{h/p}} |v_i(T_p)| \right)^{1/2} \left(\sum_{T_p \in \mathcal{T}_i^{h/p}} |u|_{S+1, v_i(T_p)}^2 \right)^{1/2}, \\
&\leq C (h/p)^{S+1} |V_i|^{1/2} \left(\sum_{T_p \in \mathcal{T}_i^{h/p}} |u|_{S+1, v_i(T_p)}^2 \right)^{1/2}, \tag{4.26}
\end{aligned}$$

where we have used the fact that the sub-volumes partition V_i .

To obtain the final result (4.20), substitute the local estimate (4.26) into the reaction functional (4.15) and simplify as in (3.37): focusing on the H^{S+1} semi-norm of u , we have

$$\begin{aligned}
\sum_{T_p \in \mathcal{T}_i^{h/p}} |u|_{S+1, v_i(T_p)}^2 &= \sum_{T_p \in \mathcal{T}^{h/p}} |u|_{S+1, T_p}^2 = \sum_{T \in \mathcal{T}^h} |u|_{S+1, T}^2, \\
&\equiv |u|_{S+1, \mathcal{T}^h}^2 \leq \|u\|_{S+1, \Omega}^2, \tag{4.27}
\end{aligned}$$

where we've used the nested property of the sub-volumes and the triangulations (2.5): each sub-volume is contained in a single element of $\mathcal{T}^{h/p}$ which in turn is contained in a single element of \mathcal{T}^h .

If $u \in C^0(\bar{\Omega}) \cap H^{S+1}(\mathcal{T}^h) \cap H^1(\Omega)$ for $S \in (1/2, p]$, the result holds with $|u|_{S+1, \Omega}^2$ replaced by $|u|_{S+1, \mathcal{T}^h}^2$. (see Corollary 3.1) \square

In analogy to Lemma 3.4 of Section 3.6, we develop an approximation result for piecewise constant interpolation of elements from our FVE trial space $\mathcal{W}^h = \mathcal{S}^p(\mathcal{T}^h)$. This error estimate will be used in the next section to establish reaction and advection discrete ellipticity results.

Lemma 4.3 (Constant Interpolation) Let $\eta_0(u) = u - \Pi_0^{h/p} u$ be the interpolation error for $u \in \mathcal{W}^h$ into $\mathcal{S}^0(\mathcal{V}^{h/p})$. Then we have the following estimate:

$$\mathcal{R}(\eta_0(u)) \leq C (h/p) |u|_{1, \Omega}. \tag{4.28}$$

Proof: If $u \in \mathcal{W}^h = \mathcal{S}^p(\mathcal{T}^h)$, $u \in C^0(\overline{\Omega}) \cap H^1(\Omega)$ so that volume-wise constant interpolation and the error estimate (4.28) are well-defined. Recall that for polynomials that $\Pi_0^{h/p} u = \bar{u} = \sum_{i \in I} u_i \chi_i$ from (3.52) of Section 3.7. Therefore, we will use a Bramble-Hilbert argument similar to that used in Theorem 4.2 on each volume V_i in $\mathcal{V}^{h/p}$ and over the polynomial space \mathcal{W}^h .

First, we affine map from V_i to \widehat{V}_i and transform the reaction functional

$$|r_i(\eta_0)| = \frac{|V_i|}{|\widehat{V}_i|} |\widehat{r}_i(\widehat{\eta}_0)|. \quad (4.29)$$

Second, we bound \widehat{r}_i and find that the Bramble-Hilbert lemma holds when the regularity parameter for \widehat{u} equals 1: i.e., constant interpolation is exact for polynomials of degree 0. That is,

$$|\widehat{r}_i(\widehat{\eta}_0)| \leq C |\widehat{u}|_{1, \widehat{V}_i}. \quad (4.30)$$

Third, we transform back to each volume V_i in $\mathcal{V}^{h/p}$ and find estimates for $r_i(\eta_0)$:

$$|r_i(\eta_0)| \leq C (h/p) |V_i|^{1/2} |u|_{1, V_i}, \quad (4.31)$$

where $O(h/p)$ is the diameter of V_i . Note that (4.31) is an $L^2(\Omega)$ -equivalent optimal-order local estimate on the volume V_i .

To obtain the final result (4.28), substitute the local estimate (4.31) into the reaction functional (4.15) and simplify. \square

4.4 Ellipticity

As in Section 3.7, we derive discrete ellipticity results for the respective reaction and advection bilinear forms (4.4) and (4.2) in terms of the discrete L^2 and H^1 norms defined in Section 3.3. As we saw in the error analysis of Section 3.8 and will see in the error analysis of the next section, ellipticity is crucial. Once ellipticity is established, FVE analysis reduces to the approximation theory developed in Sections 3.6 and 4.3.

As in FE, our FVE ellipticity results allow us to consider two types of general elliptic equations: steady reaction-diffusion and advection-reaction-diffusion with an “artificial” reaction coefficient that allows us to overcome the effects of the non-symmetric advection term. We note that in the latter case, the velocity field is general. We cannot consider a steady advection-diffusion equation without an “artificial” velocity field: $\nabla \cdot \mathbf{a} \geq \epsilon > 0$ as in [33]. In the context of parabolic equations, we will see in Chapter 5 that artificial reaction is only needed to define an elliptic projection (discussed in Section 4.6) of the exact parabolic solution. After the projection is established, we can consider transient reaction-diffusion, advection-diffusion, and full-blown advection-reaction-diffusion equation with general “physical” coefficients. Therefore, the full fruit of our approach for elliptic equations will be realized in the more general context of parabolic equations.

4.4.1 Reaction

For reaction we find an ellipticity result analogous to the diffusion result (3.53). Using the approximation result (4.28) of Lemma 4.3 from the previous section, we demonstrate an “asymptotic” ellipticity result in terms of

our discrete FVE trial space $\mathcal{S}^p(\mathcal{T}^h)$ and our discrete L^2 norm for the reaction bilinear form. That is, for (h/p) sufficiently small the ellipticity constant μ is independent of the computational mesh. Again this type of argument is necessary to maintain the FE formalism in our FVE analysis, since our discrete L^2 is defined to be the restriction of the continuous L^2 norm to piecewise constant functions on $\mathcal{V}^{h/p}$.

Lemma 4.4 There exists a positive constant $\mu = \mu(r_m; \lambda_r)$, for $h/p \leq \lambda_r$, such that

$$\mathcal{R}(u, \bar{u}) \geq \mu |u|_{0,\omega}^2 \quad \forall u \in \mathcal{W}^h, \quad (4.32)$$

where r_m is the positive lower bound on the reaction coefficient, and $\lambda_r = \frac{1}{2} r_m C^{-1} |u|_{0,\omega} / |u|_{1,\Omega}$ is an asymptotic ellipticity parameter.

Proof: In analogy to the proof of the diffusion result, we start with a sub-result. If u in (4.4) is replaced by the lumped representation $\bar{u} \in \mathcal{X}^h$ of (3.52), a discrete ellipticity condition is easily demonstrated:

$$\begin{aligned} \mathcal{R}(\bar{u}, \bar{u}) &= \sum_{i \in I} u_i^2 \int_{V_i} r \, d\mathbf{x}, \\ &\geq r_m |u|_{0,\omega}^2. \end{aligned} \quad (4.33)$$

For the general reaction term (4.4), recall that $\bar{u} \equiv \Pi_0^{h/p} u$. Now, we analyze the splitting induced by \bar{u} :

$$\begin{aligned} \mathcal{R}(u, \bar{u}) &= \mathcal{R}(\bar{u}, \bar{u}) + \mathcal{R}(u - \bar{u}, \bar{u}), \\ &\geq r_m |u|_{0,\omega}^2 - \mathcal{R}(\eta_0(u)) |u|_{0,\omega}, \\ &\geq r_m |u|_{0,\omega}^2 - C (h/p) |u|_{1,\Omega} |u|_{0,\omega}, \\ &\geq \frac{1}{2} r_m |u|_{0,\omega}^2, \end{aligned} \quad (4.34)$$

where $h/p \leq \lambda_r = \frac{1}{2} r_m C^{-1} |u|_{0,\omega} / |u|_{1,\Omega}$.

In the first three lines of (4.34), we've used the sub-result (4.33), the bound (4.14), and the approximation result (4.28), respectively. Thus, the lemma is proved with $\mu = \frac{1}{2} r_m$. \square

Therefore, (4.32) and (3.53) imply discrete ellipticity and will imply optimal-order and superconvergence error estimates for steady reaction-diffusion equations when h/p is sufficiently small.

4.4.2 Advection

For advection, our asymptotic discrete ellipticity result must be used in conjunction with the corresponding diffusion (3.53) and reaction (4.32) results and not in isolation.

Lemma 4.5 For $(h/p) \leq \min\{\lambda_{\mathbf{div}}, \lambda_{\mathbf{a}}\}$, we have the following result:

$$\mathcal{A}(u, \bar{u}) \geq -\frac{1}{2}\beta |u|_{1,\omega}^2 - \nu |u|_{0,\omega}^2, \quad \forall u \in \mathcal{W}^h, \quad (4.35)$$

where $\beta = \frac{1}{2} D_m$ is the ellipticity constant from (3.53), D_m is the positive lower bound on the diffusion coefficient, $\nu = \nu(D_m, M_{\mathbf{div}}, M_{\mathbf{a}}; \lambda_D, \lambda_{\mathbf{div}}, \lambda_{\mathbf{a}})$ is a positive constant equal to $4 \max\{M_{\mathbf{a}}^2 \max\{D_m^{-1}, 1\}, M_{\mathbf{div}}\}$, $M_{\mathbf{a}}$ and $M_{\mathbf{div}}$ are the upper bounds on the velocity vector and its divergence, and $\lambda_{\mathbf{div}} = M_{\mathbf{div}} C^{-1} |u|_{0,\omega} / |u|_{1,\Omega}$ and $\lambda_{\mathbf{a}} = M_{\mathbf{a}} C^{-1} |u|_{0,\omega} / |u|_{2,\mathcal{T}^h}$ are asymptotic ellipticity parameters.

Proof: Using the definition of (4.2) and the Divergence Theorem, we split (4.2) into $L^2(\Omega)$ and $H^1(\Omega)$ -equivalent components:

$$\mathcal{A}(u, \bar{u}) = \sum_{i \in I} \int_{V_i} \nabla \cdot \mathbf{a} u \bar{u} \, d\mathbf{x} + \sum_{i \in I} \int_{V_i} \mathbf{a} \cdot \nabla u \bar{u} \, d\mathbf{x}. \quad (4.36)$$

If the velocity field in our elliptic problem (1.1) represents an incompressible flow, $\nabla \cdot \mathbf{a} = 0$ and we can drop the first sum on the right-hand side of (4.36). However, we consider a general velocity field with a bounded divergence ($\|\nabla \cdot \mathbf{a}\| \leq M_{\text{div}}$). To this end, we define a divergence bilinear form $\widetilde{\mathcal{R}}(u, \bar{u})$ and functional $\widetilde{\mathcal{R}}(u)$ by replacing r with $\nabla \cdot \mathbf{a}$ in the definitions of (4.4) and (4.16). We note that all results for the reaction bilinear form and functional hold for the new divergence bilinear form and functional with minor modifications.

Employing these new definitions for the divergence term, we analyze the splitting induced by \bar{u} :

$$\begin{aligned}
\sum_{i \in I} \int_{V_i} \nabla \cdot \mathbf{a} u \bar{u} \, d\mathbf{x} &= \widetilde{\mathcal{R}}(\bar{u}, \bar{u}) + \widetilde{\mathcal{R}}(\eta_0(u), \bar{u}), \\
&\geq -M_{\text{div}} |u|_{0,\omega}^2 - \widetilde{\mathcal{R}}(\eta_0) |u|_{0,\omega}, \\
&\geq -M_{\text{div}} |u|_{0,\omega}^2 - C(h/p) |u|_{1,\Omega} |u|_{0,\omega}, \\
&\geq -2M_{\text{div}} |u|_{0,\omega}^2,
\end{aligned} \tag{4.37}$$

where we have used the bound (4.14), the sub-result (4.33), and the approximation result (4.28) to obtain (4.37) when $h/p \leq \lambda_{\text{div}} = M_{\text{div}} C^{-1} |u|_{0,\omega} / |u|_{1,\Omega}$.

For the second sum on the right-hand side of (4.36), we analyze the splitting induced by $u_1 = \Pi_1^{h/p} u$ with the help of some standard approximation theory (see Ciarlet [11] Theorem 17.1 and Remark 17.2) for $p \geq 2$:

$$\begin{aligned}
\sum_{i \in I} \int_{V_i} \mathbf{a} \cdot \nabla u \bar{u} \, d\mathbf{x} &= \sum_{i \in I} \int_{V_i} \mathbf{a} \cdot \nabla u_1 \bar{u} \, d\mathbf{x} + \sum_{i \in I} \int_{V_i} \mathbf{a} \cdot \nabla (u - u_1) \bar{u} \, d\mathbf{x}, \\
&\geq -M_{\mathbf{a}} |u_1|_{1,\Omega} |\bar{u}|_{0,\Omega} - M_{\mathbf{a}} C(h/p) |u|_{2,\mathcal{T}^h} |\bar{u}|_{0,\Omega}, \\
&\geq -M_{\mathbf{a}} |u|_{1,\omega} |u|_{0,\omega} - M_{\mathbf{a}}^2 |u|_{0,\omega}^2, \\
&\geq -\frac{1}{2} \beta |u|_{1,\omega}^2 - M_{\mathbf{a}}^2 \left(1 + \frac{1}{2\beta}\right) |u|_{0,\omega}^2,
\end{aligned}$$

$$\geq -\frac{1}{2}\beta|u|_{1,\omega}^2 - 2M_{\mathbf{a}}^2 \max\{D_m^{-1}, 1\} |u|_{0,\omega}^2, \quad (4.38)$$

where we have used the Cauchy-Schwarz inequality, standard H^1 approximation theory, the definitions of the L^2 and H^1 discrete norms (3.12) and (3.14), and the fact that $\beta = \frac{1}{2} D_m$ in the proof of (3.53) for $p \geq 2$ to obtain (4.38) for $h/p \leq \lambda_{\mathbf{a}} = M_{\mathbf{a}} C^{-1} |u|_{0,\omega} / |u|_{2,\mathcal{T}^h}$. For $p = 1$, the splitting is not needed; however, (4.38) is still an acceptable lower bound without requiring that (h/p) be sufficiently small.

Combining the results of (4.37) and (4.38), we obtain the desired result:

$$\begin{aligned} \mathcal{A}(u, \bar{u}) &\geq -\frac{1}{2}\beta|u|_{1,\omega}^2 - 2(M_{\mathbf{a}}^2 \max\{D_m^{-1}, 1\} + M_{\mathbf{div}}) |u|_{0,\omega}^2, \\ &\geq -\frac{1}{2}\beta|u|_{1,\omega}^2 - 4 \max\{M_{\mathbf{a}}^2 \max\{D_m^{-1}, 1\}, M_{\mathbf{div}}\} |u|_{0,\omega}^2, \end{aligned} \quad (4.39)$$

where the constant in front of $|u|_{0,\omega}^2$ is defined as ν . \square

Then (4.35), with (3.53) and (4.32), implies discrete ellipticity for steady advection-reaction-diffusion equations and will imply optimal-order and superconvergence error estimates when h/p is sufficiently small and $\mu = \mu(D_m, r_m, M_{\mathbf{div}}, M_{\mathbf{a}}; \lambda_D, \lambda_r, \lambda_{\mathbf{div}}, \lambda_{\mathbf{a}})$ is sufficiently large—this will be quantified in the next section.

4.5 Error Analysis

With the results of the previous sections, the basic structure of the FVE error analysis (3.60) presented in Section 3.8 is relatively unchanged. We will outline modifications for steady reaction-diffusion and advection-reaction-diffusion equation separately.

As before, we have employed the error splitting: $u - u^h = \zeta - \eta_p(u)$, where $u \in \mathcal{W}^+$ and $u^h \in \mathcal{W}^h$ are the exact and FVE solutions, respectively, $\zeta = \Pi_p^h u - u^h \in \mathcal{W}^h$ is the trial space representation error, and $\eta_p(u) = \Pi_p^h u - u \in \mathcal{H}$ is the trial space interpolation error.

4.5.1 Reaction-Diffusion

Combining the ellipticity results (3.53) and (4.32) of Sections 3.7 and 4.4, respectively, we find that the first line of (3.60) is modified to read

$$\beta |\zeta|_{1,\omega}^2 \leq \mathcal{D}(\zeta, \bar{\zeta}) + \mathcal{R}(\zeta, \bar{\zeta}). \quad (4.40)$$

Proof: For (h/p) sufficiently small, we have

$$\begin{aligned} \mathcal{D}(\zeta, \bar{\zeta}) + \mathcal{R}(\zeta, \bar{\zeta}) &\geq \beta |\zeta|_{1,\omega}^2 + \mu |u|_{0,\omega}^2, \\ &\geq \beta |\zeta|_{1,\omega}^2. \end{aligned} \quad (4.41)$$

In the second line of (3.60), the zero property is unchanged: subtracting (4.7) from (4.5) yields $B(u - u^h, w) = 0, \forall w \in \mathcal{X}^h$. Then, in terms of the error splitting, we have

$$\mathcal{D}(\zeta, \bar{\zeta}) + \mathcal{R}(\zeta, \bar{\zeta}) = \mathcal{D}(\eta_p, \bar{\zeta}) + \mathcal{R}(\eta_p, \bar{\zeta}). \quad (4.42)$$

Combining the upper bound results (3.17) and (4.14) of Sections 3.4

and 4.2, the third line of (3.60) becomes

$$\mathcal{D}(\eta_p, \bar{\zeta}) + \mathcal{R}(\eta_p, \bar{\zeta}) \leq \mathcal{D}(\eta_p)|\zeta|_{1,\omega} + \mathcal{R}(\eta_p)|\zeta|_{0,\omega}. \quad (4.43)$$

4.5.2 Advection-Reaction-Diffusion

Combining the ellipticity results (3.53), (4.32), and (4.35) of Sections 3.7 and 4.4, respectively, we find that the first line of (3.60) is modified to read

$$\frac{1}{2}\beta|\zeta|_{1,\omega}^2 \leq \mathcal{A}(\zeta, \bar{\zeta}) + \mathcal{D}(\zeta, \bar{\zeta}) + \mathcal{R}(\zeta, \bar{\zeta}). \quad (4.44)$$

Proof: For (h/p) sufficiently small, we have

$$\begin{aligned} \mathcal{A}(\zeta, \bar{\zeta}) + \mathcal{D}(\zeta, \bar{\zeta}) + \mathcal{R}(\zeta, \bar{\zeta}) &\geq \beta\left(1 - \frac{1}{2}\right)|\zeta|_{1,\omega}^2 + (\mu - \nu)|u|_{0,\omega}^2, \\ &\geq \frac{1}{2}\beta|\zeta|_{1,\omega}^2, \end{aligned} \quad (4.45)$$

where we have obtained (4.45) for $\mu = \frac{1}{2}r_m$ sufficiently large: $r_m \geq 2\nu = 8 \max\{M_a^2 \max\{D_m^{-1}, 1\}, M_{\text{div}}\}$. For example, if the velocity field is divergence-free and our elliptic equation (1.1) is nearly hyperbolic in character ($M_{\text{div}} = 0$ and $D_m \ll 1$), we require $r_m \geq 8M_a^2/D_m$.

In the second line of (3.60), the zero property is unchanged: subtracting (4.7) from (4.5) yields $B(u - u^h, w) = 0, \forall w \in \mathcal{X}^h$. Then in terms of the error splitting, we have

$$\mathcal{A}(\zeta, \bar{\zeta}) + \mathcal{D}(\zeta, \bar{\zeta}) + \mathcal{R}(\zeta, \bar{\zeta}) = \mathcal{A}(\eta_p, \bar{\zeta}) + \mathcal{D}(\eta_p, \bar{\zeta}) + \mathcal{R}(\eta_p, \bar{\zeta}). \quad (4.46)$$

Combining the upper bound results (3.17), (4.9), and (4.14) of Sections 3.4 and 4.2, the third line of (3.60) becomes

$$B(\eta_p(u), \bar{\zeta}) \leq (\mathcal{A}(\eta_p) + \mathcal{D}(\eta_p))|\zeta|_{1,\omega} + \mathcal{R}(\eta_p)|\zeta|_{0,\omega}. \quad (4.47)$$

4.5.3 General Error Estimate

With the modifications of the last two subsections, we apply the discrete Poincaré inequality (3.16) of Section 3.3 to $|\zeta|_{0,\omega}$ in (4.43) and (4.47). Then we combine the $H^1(\Omega)$ -equivalent optimal-order and superconvergence approximation results (3.30) and (3.41) of Section 3.6 with the $L^2(\Omega)$ -equivalent optimal-order approximation results (4.18) and (4.20) of Section 4.3. As a result, we find that FVE error analysis reduces to approximation theory. \square

Theorem 4.3 Under the conditions stated above, the optimal-order and superconvergence error estimate of (3.61) holds for the FVE approximation to the exact solution of steady reaction-diffusion and advection-reaction-diffusion equations defined by the general elliptic problem (1.1). That is,

$$|u - u^h|_{1,\omega} \leq C (h/p)^S \|u\|_{S+1,\Omega}, \quad S \in (1/2, \rho], \quad (4.48)$$

where $\rho \equiv \max\{p, P\}$, p is the order of the polynomial trial space, P is the superconvergence (for $p = 1$) parameter of Theorem 3.2: P is 1 or 2 in the absence or presence, respectively, of volume symmetry.

In our treatment of advection, we have employed a standard approach to establish a baseline analysis. In practice, we would be well-served to employ “upwinding” procedures such as those discussed in Hughes [26], Lazarov *et al.* [32, 33], and Roache [43]. In a forthcoming companion paper, upwind FVE are defined for linear finite elements and are analyzed: once again, optimal-order and superconvergence results are obtained.

We conclude our presentation for elliptic equations in the next section with a discussion of the elliptic projection operator that will be used in the FVE analysis for parabolic equations in the next chapter.

4.6 Elliptic Projection

For a general function $g \in \mathcal{W}^+ \equiv H^{S+1}(\Omega) \cap H_0^1(\Omega)$, $S \in (1/2, \rho]$, we define an elliptic map $\tilde{g}^h \in \mathcal{W}^h$ that satisfies a zero property with respect to the FVE B -form:

$$B(\tilde{g}^h - g, w) = 0, \quad \forall w \in \mathcal{X}^h. \quad (4.49)$$

Since (4.49) is only a restatement of the zero property from the previous section, Theorem 4.3 immediately yields an error estimate for the elliptic map. \square

Lemma 4.6

$$|(\tilde{g}^h - g)|_{1,\omega} \leq C (h/p)^S \|g\|_{S+1,\Omega}, \quad S \in (1/2, \rho]. \quad (4.50)$$

Notice that (4.49) actually defines a (subspace) projection \mathcal{P}_B from $\mathcal{H} \equiv \mathcal{W}^+ \oplus \mathcal{W}^h$ into $\mathcal{W}^h \subset \mathcal{H}$:

$$\mathcal{P}_B : g \in \mathcal{H} \rightarrow \tilde{g}^h \in \mathcal{W}^h \subset \mathcal{H}; \quad (4.51)$$

$$\mathcal{P}_B(\tilde{g}^h) \equiv \tilde{g}^h. \quad (4.52)$$

More generally, let $g = g(t)$ be a continuous function of time ($t \in \Theta \equiv (0, T]$); then, we have the following error estimates for continuous-time and discrete-time elliptic mappings:

Theorem 4.4 If \tilde{g}^h is continuously defined for all $t \in \Theta \equiv (0, T]$, the continuous-time elliptic mapping error satisfies:

$$\sup_{t \in \Theta} |(\tilde{g}^h - g)(\cdot, t)|_{1,\omega} \leq C (h/p)^S \|g\|_{L^\infty(H^{S+1})}, \quad S \in (1/2, \rho], \quad (4.53)$$

where $\|\cdot\|_{L^\infty(H^{S+1})}$ is the $L^\infty(\Theta)$ norm of the $H^{S+1}(\Omega)$ norm of $g(t)$:

$$\|g\|_{L^\infty(H^{S+1})} \equiv \sup_{t \in \Theta} \|g(t)\|_{S+1,\Omega}. \quad (4.54)$$

However, if \tilde{g}^h is defined only at discrete time intervals ($t \in \theta \equiv \{t^0 = 0, t^1, \dots, t^N = T\}$), the discrete-time elliptic mapping error satisfies:

$$\max_{t \in \theta} |(\tilde{g}^h - g)(\cdot, t)|_{1, \omega} \leq C (h/p)^S \|g\|_{L^\infty(H^{S+1})}, \quad S \in (1/2, \rho]. \quad (4.55)$$

Proof: To the result of Lemma 4.6, apply the definitions of sup, max, and $L^\infty(\Theta)$. \square

In the parabolic analysis of the next chapter, the elliptic mapping error will assume the role formerly played by the interpolation error. As a result, the error estimates of Theorem 4.4 will play a prominent role.

4.7 Summary

An *a priori* FVE error analysis for elliptic partial differential equations with smooth coefficients has been presented: $H^1(\Omega)$ -equivalent optimal-order and superconvergence results have been obtained in the $H^1(\omega)$ seminorm. These results can be extended to problems with discontinuous and nonlinear coefficients: the modifications for discontinuous coefficients are analogous to those of Samarskii *et al.* [46] for FV; the modifications for nonlinear coefficients are analogous to those of Douglas [12] for FE.

For the purpose of analysis, we have viewed FVE as a locally conservative Petrov-Galerkin FE and performed an FE-style error analysis. Building upon the foundation laid by Cai, Mandel, and McCormick [9] for steady diffusion equations discretized by linear finite elements in two spatial dimensions, we have extended and generalized FVE analysis to cover steady advection-reaction-diffusion equations discretized by C^0 polynomial finite elements of arbitrary order in three spatial dimensions. In addition, we have considered h , p , and h - p version FVE where increased accuracy is obtained by refining the computational mesh parametrized by h and/or by increasing the order p of the polynomial trial space.

In Sections 1.1 and 1.2 of Chapter 1, we outlined the origin of fluid dynamics and its basis on local conservation laws. Also, we motivated the need for FVE which preserves local conservation on the discrete level. In addition, we saw that FVE systematically discretizes the unknown solution, known source term data and boundary conditions on domains with complex geometries via the FE triangulation \mathcal{T}^h , the FVE triangulation $\mathcal{T}^{h/p}$, and the

FVE volumization $\mathcal{V}^{h/p}$, as discussed in Sections 2.1 and 2.2.

Motivated by the success of h , p , and h - p version FE in the structural mechanics literature (see [2, 51]), we defined h , p , and h - p version FVE for fluid dynamics problems in Section 2.3. Given the documented [39] success of Multi-level solution algorithms for h -version FVE vs. FE discretizations and the documented [25, 35, 37, 42] success of similar Multi-level and Domain Decomposition solution algorithms for p and h - p version FE discretizations, we expect similar implementation efficiency successes for p and h - p version FVE discretizations.

Furthermore, we saw that the computational set-up costs for FVE were less expensive than the costs for FE in Section 2.4. That is, if d is the spatial dimension of the problem domain, the quadrature rules used to evaluate FVE integrals and set-up a system of linear equations were seen to be $O(2^d)$ less expensive than corresponding FE quadrature rules: the quadrature operation count for reaction and source term integrals and is $O(p)^d$ for FVE and $O(2p)^d$ for FE. However, for reaction and source terms, FVE integrals are defined on volumes from $\mathcal{V}^{h/p}$ with diameters of $O(h/p)$ whereas FE integrals are defined on elements from \mathcal{T}^h with diameters of $O(h)$; therefore, the FVE quadratures are performed on regions that are $O(p)^d$ smaller than their FE counterparts. This feature makes the FVE equations more local in character and the FVE mass matrix arising from the reaction term more strongly diagonally dominant than its FE counterpart; from the standpoint of numerical linear algebra and implementation, this means that it should be easier to precondition and iteratively solve FVE systems of linear equations than their FE counterparts. In

addition, FVE diffusive and advective fluxes are defined as surface integrals which are simpler (i.e., of lower dimension) than the corresponding volume integrals for FE: the quadrature operation count is $O(p)^{d-1}$ for FVE and $O(2p)^d$ for FE—a increased savings of $O(p2^d)$.

Therefore, we expected FVE to outperform FE in standard measures of computational efficiency. Furthermore, from the standpoint of the *a priori* error analysis presented in the Chapters 3 and 4, we saw that the accuracy of FVE approximations to the exact solution of general elliptic equations either matched or surpassed FE asymptotic convergence rates for h , p , and h - p versions of the respective methods.

After introducing a Sobolev space infrastructure and a standard h , p , and h - p version FE analysis in Chapter 2, we detailed an h , p , and h - p version FVE analysis for diffusion equations in Chapter 3.

In Sections 3.1 and 3.2, we motivated and laid the groundwork for an FE-style analysis for FVE based on the formalisms introduced in the FE analysis presented in the previous chapter. After introducing discrete L^2 and H^1 Sobolev norms in Section 3.3, we demonstrated an upper bound for the FVE B -form in terms of our discrete H^1 semi-norm and an FVE diffusion functional in Section 3.4. After introducing the Bramble-Hilbert lemma in Section 3.5, we focused on an $H^1(\Omega)$ -equivalent h , p , and h - p version optimal-order approximation theory and an h -version superconvergent approximation theory for our FVE diffusion functional in Section 3.6. After establishing an optimal-order result in Theorem 3.1 for interpolation by C^0 piecewise polynomials of arbitrary

order on a regular triangulation and volumization, we focused on superconvergence for linear interpolation on a “symmetric” volumization in Theorem 3.2. That is, from the perspective of the FVE diffusion functional, the approximating power of a linear trial space on a symmetric volumization matches the power of a quadratic trial space on a regular volumization. At the core of the superconvergence theory for the diffusive flux functional is the 1-d fact that centered differencing is exact for quadratics; we invoke certain volume symmetries, discussed below, to extend this 1-d result to multi-dimensional settings.

As we saw in Sections 3.6.2 and 3.6.2.2, the definition of volume symmetry is dimensionally and geometrically dependent. That is, if $d = 1$, volume symmetry is X -symmetry: the volume boundary γ_{ij} , a point in 1-d, is at the midpoint of each element edge X_{ij} ; in this simple case, the diffusive flux is represented by a centered difference quotient. If $d = 2$, volume symmetry for C^0 linears on triangles requires X and γ -symmetries: the volume edge γ_{ij} and the element edge X_{ij} are perpendicular bisectors of one another; however for C^0 bilinears on rectangles, volume symmetry requires only X -symmetry, since the quadratic component of a piecewise bilinear trial space obviated the γ_{ij} -symmetry requirement.

When $d = 3$, volume symmetry for C^0 linears on tetrahedra requires X -symmetry (the planar volume interface γ_{ij} is a perpendicular bisector for the element edge X_{ij}) and a “centroid” γ -symmetry: i.e., $X_{ij} - \gamma_{ij}$ and the point of intersection P_{ij} is the centroid for γ_{ij} ; furthermore, the geometry of γ_{ij} is such that the centroid quadrature rule is exact for linear functions— γ_{ij} is triangular,

rectangular, hexagonal, etc. (see [49] for details). For C^0 bilinears (i.e., $\Pi_{1,1}u \in \text{span}\{1, x, y\} \otimes \{1, z\}$) on triangular prisms formed in a tensor-product mesh, volume symmetry requires X -symmetry and a “2-d” γ -symmetry: the x and y cross-sections of γ_{ij} must possess the centroid symmetry. That is, linearity in x and y requires the 2-d γ_{ij} -symmetry in x and y , but bilinearity in z eliminates the 2-d γ_{ij} -symmetry in z . For C^0 trilinears on rectangular bricks formed by a tensor-product mesh, volume symmetry requires X -symmetry and a “1-d” γ_{ij} -symmetry: either one of the 1-d cross-sections of γ_{ij} must possess the centroid symmetry. That is, the quadratic components of a trilinear trial space obviate γ_{ij} -symmetry requirements as in 2-d; however, its cubic component invokes a reduced, 1-d γ_{ij} -symmetry.

Finally, using the “degenerate volume paradigm” of Section 3.6.2.2, we see that linear FVE on a symmetric volumization is identical to a quadratic FVE on a degenerate volumization; hence, superconvergence for linear FVE can be viewed as a special case of the optimal-order convergence for quadratic FVE.

When discussing superconvergence for linear trial spaces, it is important to emphasize that it is primarily a local phenomenon. Qualitatively, global superconvergence is the cumulative effect of local volume symmetries or uniformities in a given computational mesh. Quantitatively, local quadratic convergence leads to superlinear global convergence: the precise rate depends upon the prevalence of local volume symmetry or uniformity conditions on a global computational mesh. As a result, local quadratic convergence of a numerical solution (in regions of symmetry/uniformity on a non-uniform mesh)

“remediates” the global solution: superlinear, rather than linear, convergence is observed globally.

Using the optimal-order approximation results of Section 3.6.1, we derived (for $p \geq 2$) an “asymptotic” ellipticity result for the diffusion bilinear form in terms of our discrete FVE trial space $\mathcal{S}^p(\mathcal{T}^h)$ and our discrete H^1 semi-norm in Section 3.7. That is, for (h/p) sufficiently small, the ellipticity constant is independent of the computational mesh. This type of argument is unavoidable if we wish to maintain the FE formalism in our FVE error analysis for higher-order ($p \geq 2$) polynomial trial spaces, since our discrete H^1 semi-norm is defined as the restriction of the continuous H^1 semi-norm to piecewise linear polynomials on the FVE triangulation $\mathcal{T}^{h/p}$. Finally, in Section 3.8, we demonstrate an FE-style error analysis for h , p , and h - p version FVE that matches or exceeds the results of the standard h , p , and h - p FE analysis presented in Chapter 2.

In the present chapter, we extended the scope of our analysis to general elliptic equations and detailed the analogous theoretical results for the FVE treatment of advection and reaction. After deriving upper bounds for the FVE advection and reaction bilinear forms in terms of our discrete norms and FVE advection and reaction functionals in Section 4.2, we focused our attention on developing an approximation theory for our advection and reaction functionals. In order to avoid potential conflicts with the superconvergence result for the FVE diffusion functional, we developed an h , p , and h - p version $L^2(\Omega)$ -equivalent optimal-order approximation theory for our advection and reaction functionals in Section 4.3. Using these approximation results,

we demonstrated “asymptotic” ellipticity results in terms of our discrete FVE trial space $\mathcal{S}^p(\mathcal{T}^h)$ and our discrete L^2 (and H^1) norm(s) for the reaction and advection bilinear forms in Section 4.4. That is, for (h/p) sufficiently small the ellipticity constants are independent of the computational mesh. Again this type of argument is necessary to maintain the FE formalism in our FVE analysis, since our discrete L^2 is defined to be the restriction of the continuous L^2 norm to piecewise constant functions on $\mathcal{V}^{h/p}$. In Section 4.5, we demonstrate an FE-style error analysis for steady reaction-diffusion and advection-reaction-diffusion equations for h , p , and h - p version FVE that matches or exceeds the corresponding FE results. And in the previous section, we concluded our presentation for elliptic equations with a discussion and analysis of the FVE elliptic projection operator to set the stage for an h , p , and h - p version FVE analysis for parabolic equations.

In the next chapter, the elliptic analysis presented here will be extended to general parabolic equations in a straightforward manner by an FVE variant of the elliptic projection argument of Wheeler [55]. That is, an FE-style error analysis of a continuous-time method that uses ordinary differential equations in time and FVE in space (as in a “method of lines” procedure) will yield error estimates of the form

$$\sup_{t \in (0, T]} |(u - u^h)(\cdot, t)|_{1, \omega} \leq C (h/p)^S (\|\partial_t u\|_{L^2(H^S)} + \|u\|_{L^\infty(H^{S+1})}); \quad (4.56)$$

similarly, an FE-style error error analysis of a method that uses finite differences in time and FVE in space (as in a fully-implicit “time-marching” procedure)

will yield error estimates of the form

$$\begin{aligned}
& \max_{1 \leq n \leq N} |(u - u^h)(\cdot, t^n)|_{1,\omega} \\
& \leq C \left(\Delta t^R \|\partial_t^{R+1} u\|_{L^2(L^2)} + \right. \\
& \quad \left. (h/p)^S (\|\partial_t u\|_{L^2(H^S)} + \|u\|_{L^\infty(H^{S+1})}) \right), \quad (4.57)
\end{aligned}$$

where $S \in (1/2, \rho]$ and $R = 1$ or 2 depending on whether backward Euler ($R = 1$) or Crank-Nicolson ($R = 2$) time differencing is used.

5. Analysis for Parabolic Equations

Our FVE analysis for parabolic equations is based on the model problem posed on bounded spatial ($\Omega \subset \mathfrak{R}^d, d \in \{1, 2, 3\}$) and temporal ($\Theta \equiv (0, T]$) domains:

$$m \partial_t u + \nabla \cdot (\mathbf{a} u - D \nabla u) + r u = f \quad \text{in } \Omega \times \Theta, \quad (5.1)$$

subject to a homogeneous Dirichlet boundary condition,

$$u(\cdot, t)|_{\partial\Omega} = 0, \quad t \in \Theta$$

and an initial condition,

$$u(\mathbf{x}, 0) = u_0(\mathbf{x}), \quad \mathbf{x} \in \Omega.$$

In (5.1), m is a mass storage coefficient that is continuous, bounded, and non-degenerate ($0 < \underline{m} \leq m \leq \overline{m}$).

More to the point, (5.1) arises from an integral conservation law (in space) that holds for all $t \in \Theta$: let $u = u(t)$ and $f = f(t)$;

$$\int_V m \partial_t u \, d\mathbf{x} + \int_{\partial V} (\mathbf{a} u - D \nabla u) \cdot \mathbf{n} \, dS + \int_V r u \, d\mathbf{x} = \int_V f \, d\mathbf{x}, \quad \forall V \subseteq \Omega. \quad (5.2)$$

Following the algorithm development for elliptic equations, the finite volume element method for parabolic equations is based on (5.2) rather than (5.1). Starting from (5.2), FVE replaces $u(t)$ and $f(t)$ with suitable FE approximations (in space) on the triangulations \mathcal{T}^h and $\mathcal{T}^{h/p}$ of Ω and poses the conservation law on a finite set of volumes from a volumization $\mathcal{V}^{h/p}$ of Ω .

Following the analysis for elliptic equations, we transition from a finite number of local laws based on (5.2) to a global law (formed by a weighted summation of the local laws) that holds for all $t \in \Theta$:

$$\mathcal{M}(\partial_t u(t), w) + B(u(t), w) = (f(t), w), \quad \forall w \in \mathcal{X}^h, \quad (5.3)$$

where the mass storage bilinear form is defined as:

$$\mathcal{M}(\partial_t u(t), w) = \sum_{i \in I} \int_{V_i} m \partial_t u(t) w_i d\mathbf{x} \quad (5.4)$$

and $B(u(t), w)$ and $(f(t), w)$ are defined as in the previous chapters.

In addition to the usual spatial regularity assumption $u \in H^{S+1}(\Omega) \cap H_0^1(\Omega)$, $S \in (1/2, \rho]$ of the previous chapters, we make the temporal regularity assumption $u(t) \in H^{R+1}(\Theta)$, $R > 1/2$, so that $u(t)$ can be thought of as a continuously differentiable function of time (by the Sobolev embedding theorem).

In this chapter, we outline an FVE analysis for parabolic equations. First, we define continuous-time and discrete-time numerical methods for general parabolic equations in Section 5.1: the continuous-time method is an FVE variant of the “method of lines”; the discrete-time methods are FVE variants of Backward Euler and Crank-Nicholson “time marching” methods. In Sections 5.2 and 5.3, the components of FVE parabolic analysis are developed and assembled to yield optimal-order and superconvergence results for h , p , and h - p version continuous-time and discrete-time FVE. Future work based on this thesis is outlined in Section 5.5.

5.1 Numerical Methods

In the previous chapters, we have focused on the spatial treatment of unknown u (and known data f) in (5.2); now we turn to its temporal treatment. First, we consider leaving $u(t)$ continuous in time; second, we consider (backward Euler and Crank-Nicolson) discretizations in time.

5.1.1 Continuous-Time

In a “method of lines” or continuous-time FVE method, we follow (5.2) posed only on volumes from $\mathcal{V}^{h/p}$ and replace the unknown $u(t)$ by the semi-discretization:

$$u^h(t) = \sum_{i \in I} u_i^h(t) \phi_i(\mathbf{x}) \in \mathcal{W}^h = \mathcal{S}_0^p(\mathcal{T}^h), \quad (5.5)$$

where $u_i^h(t)$ is a continuous (in time) coefficient corresponding to the basis function ϕ_i for the trial space \mathcal{W}^h . Also, the source term $f = f(t)$ is replaced by the numerical source term $f^h = f^h(t) \in \mathcal{S}^p(\mathcal{T}^h)$ defined in a similar manner. However, we require for all $t \in \Theta$ that f^h satisfy the projection relationship,

$$(f - f^h, w) = 0, \quad \forall w \in \mathcal{S}^0(\mathcal{V}^{h/p}), \quad (5.6)$$

with respect to f to ensure that the numerical source term conserves mass.

To discretize the initial condition $u(0) = u_0$, we invoke the elliptic map (4.49): $u^h(0)$ satisfies

$$B(u^h(0) - u_0, w) = 0, \quad w \in \mathcal{X}^h = \mathcal{S}_0^0(\mathcal{V}^{h/p}). \quad (5.7)$$

As a result, the PDE (5.1) is replaced by a system of ODEs in time according to a FVE discretization in space. Assuming that the ODE system

is solved exactly in time, we focus on the errors induced by the spatial FVE discretization.

5.1.2 Discrete-Time

Now we consider fully-implicit “time-marching” or discrete-time FVE methods that employ a FD discretization of (5.2) in time in addition to the FVE discretization in space. First, we partition the time domain $\Theta \equiv (0, T]$ into N time levels $\theta = \{t^0 = 0, t^1, \dots, t^N = T\}$. For simplicity, we assume that Θ is uniformly partitioned: let $\Delta t \equiv T/N$ so that $t^n \equiv n\Delta t$ for $0 \leq n \leq N$.

In a Backward Euler “time-marching” or discrete-time FVE method, we follow (5.2) at each discrete time level $t^n \in \theta, 1 \leq n \leq N$ and replace the unknown $u^n \equiv u(t^n)$ by the discretization:

$$u_h^n = \sum_{i \in I} U_i^n \phi_i(\mathbf{x}) \in \mathcal{W}^h, \quad (5.8)$$

where U_i^n is a discrete-time coefficient corresponding to the basis function ϕ_i for the trial space $\mathcal{W}^h = \mathcal{S}_0^p(\mathcal{T}^h)$. As in the continuous-time method, the initial condition u_0 is replaced by its elliptic map \tilde{u}_h^0 defined in (5.7). Also, the source term $f^n \equiv f(t^n)$ is replaced by $f_h^n \in \mathcal{S}^p(\mathcal{T}^h)$ which satisfies the projection relationship,

$$(f^n - f_h^n, w) = 0, \quad \forall w \in \mathcal{S}^0(\mathcal{V}^{h/p}), \quad (5.9)$$

at each time level. Most importantly, the time derivative $\partial_t u^n$ is replaced by a finite difference approximation:

$$\partial_t u^n \equiv (u^n - u^{n-1})/\Delta t. \quad (5.10)$$

Taylor’s theorem (with integral remainder) tells us that the approximation

error is given by

$$\Phi_1^n \equiv \partial u^n - \partial_t u^n = \frac{1}{\Delta t} \int_{t^{n-1}}^{t^n} (t - t^{n-1}) \partial_t^2 u \, dt, \quad (5.11)$$

so we expect first-order accuracy in time for a backward Euler discretization.

In a Crank-Nicolson discretization, we approximate the unknown at the intermediate time level $t^{n-1/2} = (t^n + t^{n-1})/2$: $u(t^{n-1/2})$ is replaced by $u_h^{n-1/2} = (u_h^n + u_h^{n-1})/2$ and $f(t^{n-1/2})$ is replaced by $f_h(t^{n-1/2})$ subject to (5.6) defined at $t^{n-1/2}$. Most importantly, the time derivative $\partial_t u^{n-1/2}$ is replaced by the finite difference approximation $\partial_t u_h^n$. Taylor's theorem (with integral remainder) tells us that the approximation error is given by

$$\Phi_2^n \equiv \partial u^n - \partial_t u^{n-1/2} = \frac{1}{2\Delta t} \int_{t^{n-1}}^{t^n} (t - t^{n-1})(t^n - t) \partial_t^3 u \, dt, \quad (5.12)$$

so we expect second-order accuracy in time for a Crank-Nicolson discretization.

Now that we have defined numerical methods for (5.1), we proceed with a parabolic error analysis following the pattern previously established for elliptic equations as closely as possible. More precisely, we will follow an FVE variant of the standard “energy” arguments for parabolic equations developed for FV and FE by Douglas [12], Douglas and Dupont [13, 14], Ewing *et al.* [17], and Wheeler [55].

5.2 Components of Analysis

Following the analysis for elliptic equations, we transition from a finite number of local laws based on (5.2) to a global law (formed by a weighted summation of the local laws) that holds for all $t \in \Theta$:

$$\mathcal{M}(\partial_t u(t), w) + B(u(t), w) = (f(t), w), \quad \forall w \in \mathcal{X}^h, \quad (5.13)$$

where the M -form (mass storage bilinear form) is defined as:

$$\mathcal{M}(\partial_t u(t), w) = \sum_{i \in I} \int_{V_i} m \partial_t u(t) w_i dx \quad (5.14)$$

and $B(u(t), w)$ and $(f(t), w)$ are defined as in the previous chapters. Note that the form of (5.13)—the elliptic B -form is split off from the time derivative—suggests splitting the parabolic error into elliptic and non-elliptic components as in the “elliptic projection” argument of Wheeler [55].

Recalling the development of our standard elliptic analysis, we see that boundedness, approximation, and ellipticity results for the B -form can be applied to (5.13). Furthermore, similar results can be derived for the \mathcal{M} -form (5.14): the time-derivative term can be treated like the reaction term in Chapter 4 as demonstrated below. In our results, we begin with a discussion of the continuous-time result followed with the corresponding or analogous discrete-time result for (5.14) with $\partial_t u(t)$ replaced by ∂u^n .

5.2.1 Boundedness

Following the proof of Lemma 4.2 for the reaction bilinear form (4.4), the Cauchy-Schwarz inequality gives us an upper bound for the mass storage bilinear form (5.14) for continuous-time FVE in terms of the admissible trial

space $\mathcal{H} = \mathcal{W}^+ \oplus \mathcal{W}^h$. \square

Lemma 5.1

$$|\mathcal{M}(\partial_t u, w)| \leq \mathcal{M}(\partial_t u) |w|_{0,\omega}, \quad \forall u \in \mathcal{H}, w \in \mathcal{X}^h. \quad (5.15)$$

$\mathcal{M}(\cdot)$ is a bounded functional:

$$\mathcal{M}(\partial_t u) \equiv \left(\sum_{i \in I} \frac{1}{|V_i|} m_i^2(\partial_t u) \right)^{1/2}; \quad (5.16)$$

$m_i(\cdot)$ is a bounded linear functional arising from the FVE treatment of mass storage:

$$m_i(\partial_t u) = \int_{V_i} m \partial_t u \, d\mathbf{x}. \quad (5.17)$$

Similarly, we have the following upper bound for discrete-time FVE:

Lemma 5.2

$$|\mathcal{M}(\partial u^n, w)| \leq \mathcal{M}(\partial u^n) |w|_{0,\omega}, \quad \forall u \in \mathcal{H}, w \in \mathcal{X}^h. \quad (5.18)$$

With an upper bound for the mass storage bilinear form in terms of the non-standard, problem-dependent, bounded linear functional (5.17), we need to prove corresponding approximation theorems.

5.2.2 Approximation Theory

For parabolic equations, we need both spatial and temporal approximation theorems for the mass storage functional (5.16). Since we will follow the elliptic projection argument of Wheeler [55], the elliptic map will assume the role played by the interpolation map in the previous chapters.

5.2.2.1 Spatial

Since the elliptic map is idempotent (if $u \in \mathcal{S}^k(\mathcal{T}^h)$, for $1 \leq k \leq p$, $\mathcal{P}_B(u) = u$), we can apply the Bramble-Hilbert arguments developed for the interpolation map in the previous chapters with minor modifications.

Following the proof of Theorem 4.2 for the reaction functional (4.15) and proof of Theorem 4.4 for the elliptic map, the Bramble-Hilbert lemma result for the mass storage functional (5.16) for continuous-time FVE is equivalent to an optimal-order $L^2(\Omega)$ estimate with “reduced” interpolation. \square

Theorem 5.1 Let $\eta = \eta(t) = (\tilde{u}^h - u)(\cdot, t)$ for $t \in \Theta = (0, T]$ be the continuous-time elliptic mapping error. Then we have the following estimate:

$$\mathcal{M}(\partial_t \eta) \leq C (h/p)^S \|\partial_t u(t)\|_{S, \Omega}, \quad S \in (1/2, \rho]. \quad (5.19)$$

Although an order $p + 1$ result like (4.20) holds here, it is useless—the final parabolic error estimate is only order ρ . Therefore, the purpose of (5.19) is to minimize spatial regularity requirements for $\partial_t u$.

The proof of Theorem 5.1 is modified to yield the analogous result for discrete-time FVE:

Theorem 5.2 Let $J^n = [t^{n-1}, t^n]$ and $\eta^n = \eta(t^n) = (\tilde{u}^h - u)(\cdot, t^n)$, for $1 \leq n \leq N$, be the discrete-time elliptic mapping error. Then we have the following estimate:

$$\mathcal{M}(\partial \eta^n) \leq C \Delta t^{-1/2} (h/p)^S \|\partial_t u\|_{L^2[J^n; H^S(\Omega)]}, \quad S \in (1/2, \rho]. \quad (5.20)$$

where $\|\cdot\|_{L^2[J^n; H^S(\Omega)]}$ is the $L^2(J^n)$ norm of the $H^S(\Omega)$ norm of $u(t)$:

$$\|u\|_{L^2[J^n; H^S(\Omega)]} = \left(\int_{t^{n-1}}^{t^n} \|u(t)\|_{S, \Omega}^2 dt \right)^{1/2}. \quad (5.21)$$

Proof: The result follows easily upon noting that

$$\partial\eta^n = \frac{1}{\Delta t} \int_{t^{n-1}}^{t^n} \partial_t \eta \, dt; \quad (5.22)$$

then apply Fubini's Theorem and the result of the previous theorem. \square

5.2.2.2 Temporal

For the discrete-time methods, a temporal approximation theorem is needed to estimate the finite difference error for the time derivative. Recalling the (Taylor series) approximation error formulas from the previous section, we present the following result.

Theorem 5.3 Let Φ_R^n be the finite difference error of (5.11) and (5.12) for the Backward Euler ($R = 1$) and Crank-Nicolson ($R = 2$) time discretizations, then we have the following temporal error estimate:

$$\mathcal{M}(\Phi_R^n) \leq C \Delta t^{(R-1)/2} \|\partial_t^{R+1} u\|_{L^2[J^n; L^2(\Omega)]}, \quad (5.23)$$

Proof: The result (5.23) is built up from the local error estimate:

$$m_i(\Phi_R^n) \leq C \Delta t^{R-1/2} |V_i|^{1/2} \|\partial_t^{R+1} u\|_{L^2[J^n; L^2(V_i)]}, \quad (5.24)$$

which is easily derived with the aid of Fubini's Theorem and the Cauchy-Schwarz inequality. \square

5.2.3 Ellipticity

In order to use the standard “energy” argument used in the analysis of parabolic equations, we require both spatial and temporal ellipticity results.

5.2.3.1 Spatial

Following the proof of Lemma 4.4 for the reaction bilinear form, we find that the mass storage bilinear form satisfies continuous-time and discrete-time “asymptotic” ellipticity conditions. \square

Lemma 5.3 For (h/p) sufficiently small, there exists $\alpha > 0$ such that

$$\mathcal{M}(\partial_t u, \partial_t \bar{u}) \geq \alpha |\partial_t u|_{0,\omega}^2, \quad \forall u \in \mathcal{W}^h; \quad (5.25)$$

and

$$\mathcal{M}(\partial u^n, \partial \bar{u}^n) \geq \alpha |\partial u^n|_{0,\omega}^2, \quad \forall u \in \mathcal{W}^h. \quad (5.26)$$

5.2.3.2 Temporal

For the B -form, we need “energy” inequalities to bound the time evolution of continuous-time and discrete-time solutions.

Lemma 5.4 For (h/p) sufficiently small, there exists $C > 0$ such that

$$B(u, \partial_t \bar{u}) \geq \frac{1}{2} \frac{d}{dt} B(u, \bar{u}) - \frac{1}{4} \alpha |\partial_t u|_{0,\omega}^2 - C B(u, \bar{u}), \quad \forall u \in \mathcal{W}^h. \quad (5.27)$$

and

$$B(u^n, \partial \bar{u}^n) \geq \frac{1}{2} \partial B(u^n, \bar{u}^n) - \frac{1}{4} \alpha |\partial u^n|_{0,\omega}^2 - C B(u^n, \bar{u}^n), \quad \forall u \in \mathcal{W}^h, \quad (5.28)$$

where

$$\partial B(u^n, \bar{u}^n) \equiv [B(u^n, \bar{u}^n) - B(u^{n-1}, \bar{u}^{n-1})] / \Delta t. \quad (5.29)$$

Proof: We follow the outline of the proof of Lemma 4.5 for the advection bilinear form. To derive (5.27), note that

$$B(\bar{u}, \partial_t \bar{u}) = \frac{1}{2} \frac{d}{dt} B(\bar{u}, \bar{u}), \quad \forall u \in \mathcal{W}^h. \quad (5.30)$$

since $u_i \partial_t u_i = \frac{1}{2} \frac{d}{dt} u_i^2$. Then we use our approximation theory and ellipticity results for the B -form to derive a lower bound for (h/p) sufficiently small,

$$B(u - \bar{u}, \partial_t \bar{u}) \geq \frac{1}{2} \frac{d}{dt} B(u - \bar{u}, \bar{u}) - \frac{1}{4} \alpha |\partial_t u|_{0,\omega}^2 - C B(u, \bar{u}), \quad \forall u \in \mathcal{W}^h, \quad (5.31)$$

using the techniques developed in the demonstration of the advection ellipticity result (4.35).

To derive the discrete-time result (5.28), we follow the same pattern discussed above with a minor adjustment:

$$B(\bar{u}^n, \partial \bar{u}^n) \geq \frac{1}{2} \partial B(\bar{u}^n, \bar{u}^n), \quad \forall u \in \mathcal{W}^h. \quad (5.32)$$

since $(u_i^n, u_i^n - u_i^{n-1}) \geq \frac{1}{2} [(u_i^n, u_i^n) - (u_i^{n-1}, u_i^{n-1})]$ by use of the “trivial” inequality $ab \leq \frac{1}{2}(a^2 + b^2)$. \square

5.3 Error Analysis

In this section, we combine the results of the previous section to derive optimal-order and superconvergence results for h , p , and h - p version continuous-time and discrete-time FVE for the general parabolic equation (5.1). The pattern we follow closely approximates the pattern previously established for elliptic equations. But more precisely, we will follow an FVE variant of the standard “energy” arguments for parabolic equations developed for FV and FE by Douglas [12], Douglas and Dupont [13, 14], Ewing *et al.* [17], and Wheeler [55].

5.3.1 Error Splitting

In analogy to our elliptic analysis, we employ an error splitting with respect to the elliptic map $\tilde{u}^h(t)$ for the exact solution $u(t)$ to the parabolic equation (5.1): for the continuous-time method, the error splitting is $u - u^h = \zeta(t) - \eta(t)$, where $\zeta(t) \equiv (\tilde{u}_h - u^h)(\cdot, t) \in \mathcal{W}^h$ and $\eta(t) \equiv (\tilde{u}_h - u)(\cdot, t)$; for the discrete-time methods, we replace t with t^n in the error splittings. Since $|u - u^h|_{1,\omega} \leq |\zeta|_{1,\omega} + |\eta|_{1,\omega}$ and by Theorem 4.4 we have the estimates

$$\sup_{t \in (0, T]} |\eta(\cdot, t)|_{1,\omega} \leq C (h/p)^S \|u\|_{L^\infty(H^{S+1})}, \quad S \in (1/2, \rho], \quad (5.33)$$

and

$$\max_{1 \leq n \leq N} |\eta(\cdot, t^n)|_{1,\omega} \leq C (h/p)^S \|u\|_{L^\infty(H^{S+1})}, \quad S \in (1/2, \rho], \quad (5.34)$$

finding a error estimate for continuous-time and discrete-time FVE reduces to estimating $|\zeta|_{1,\omega}$.

5.3.2 Zero Property

In the discussion below, we simplify notation: let $\zeta = \zeta(t)$, for $t \in (0, T]$, and $\zeta^n = \zeta(t^n)$, for $1 \leq n \leq N$. For continuous-time FVE, we see that the parabolic error satisfies $\mathcal{M}(\partial_t(u - u^h), w) + B(u - u^h, w) = 0$, while the elliptic mapping error satisfies $B(\eta, w) = 0$. Combining these results in terms of the error splitting, we find the continuous-time zero property:

$$\mathcal{M}(\partial_t \zeta, w) + B(\zeta, w) = \mathcal{M}(\partial_t \eta, w), \quad \forall w \in \mathcal{X}^h. \quad (5.35)$$

Similarly, for the discrete-time methods, we include the temporal approximation errors Φ_R^n , for $R \in \{1, 2\}$, of (5.11) and (5.12). Therefore, the discrete-time zero property is

$$\mathcal{M}(\partial \zeta^n, w) + B(\zeta^n, w) = \mathcal{M}(\Phi_R^n, w) + \mathcal{M}(\partial \eta^n, w), \quad \forall w \in \mathcal{X}^h, \quad (5.36)$$

for the Backward Euler ($R = 1$) and Crank-Nicolson ($R = 2$) time discretizations.

5.3.3 Energy Argument

For continuous-time FVE, take $w = \partial_t \bar{\zeta}(\tau) \in \mathcal{X}^h$, for $\tau \in (0, t]$, as a test function in (5.35); then use the upper bound (5.15), the spatial and temporal ellipticity conditions (5.25) and (5.27), and some algebra to obtain:

$$\frac{1}{2} \alpha |\partial_t \zeta|_{0, \omega}^2 + \frac{1}{2} \frac{d}{dt} B(\zeta(\tau), \bar{\zeta}) \leq C [B(\zeta(\tau), \bar{\zeta}) + \mathcal{M}^2(\partial_t \eta)]. \quad (5.37)$$

For discrete-time FVE, take $w = \partial_t \bar{\zeta}(t^k) \in \mathcal{X}^h$ for $1 \leq k \leq n$ as a test function in (5.36); then the analogous steps lead to a similar result:

$$\frac{1}{2} \alpha |\partial \zeta^k|_{0, \omega}^2 + \frac{1}{2} \partial B(\zeta^k, \bar{\zeta}^k) \leq C [B(\zeta^k, \bar{\zeta}^k) + \mathcal{M}^2(\Phi_R^k) + \mathcal{M}^2(\partial \eta^k)]. \quad (5.38)$$

For continuous-time FVE, apply the continuous Gronwall inequality for differential equations to (5.37): multiply (5.37) by an integrating factor and integrate the result over $[0, t]$; note that $\zeta(0) = 0$, since the numerical initial condition is the elliptic map of the exact initial condition—recall (5.7); then apply the spatial approximation estimate (5.19) to obtain the result:

$$B(\zeta(t), \bar{\zeta}) \leq [C (h/p)^S \|\partial_t u\|_{L^2(H^S)}]^2. \quad (5.39)$$

For discrete-time FVE, apply the discrete Gronwall inequality for difference equations to (5.38): multiply (5.38) by Δt and a summation factor (which is positive for Δt sufficiently small) and sum the result from $k = 1$ to $k = n$; note that $\zeta^0 = 0$; then apply the spatial and temporal approximation estimates (5.20) and (5.23) to obtain the result:

$$B(\zeta^n, \bar{\zeta}^n) \leq C [\Delta t^R \|\partial_t^{R+1} u\|_{L^2(L^2)} + (h/p)^S \|\partial_t u\|_{L^2(H^S)}]^2. \quad (5.40)$$

Now apply the ellipticity condition for the B -form (4.45) to (5.39) and (5.40): e.g., $B(\eta(t), \bar{\eta}) \geq \frac{1}{2}\beta |\eta(t)|_{1,\omega}^2$ for the continuous-time FVE solution of the advection-reaction-diffusion equation. As a result, we obtain the following estimates for $|\zeta(t)|_{1,\omega}$ and $|\zeta^n|_{1,\omega}$:

$$\sup_{t \in (0, T]} |\zeta(\cdot, t)|_{1,\omega} \leq C (h/p)^S \|\partial_t u\|_{L^2(H^S)}, \quad S \in (1/2, \rho]. \quad (5.41)$$

and

$$\max_{1 \leq n \leq N} |\zeta(\cdot, t^n)|_{1,\omega} \leq C [\Delta t^R \|\partial_t^{R+1} u\|_{L^2(L^2)} + (h/p)^S \|\partial_t u\|_{L^2(H^S)}]. \quad (5.42)$$

Finally, we combine the previously derived estimates (5.33) and (5.34) for the elliptic map with the newly derived estimates (5.41) and (5.42) to obtain

convergence theorems for h , p , and h - p version continuous-time and discrete-time FVE numerical solution of the advection-reaction-diffusion parabolic equation (5.1). \square

Theorem 5.4 Our analysis of the continuous-time method defined in Section 5.1.1 that uses ordinary differential equations in time and FVE in space (as in “method of lines” procedure) yields the following error estimate:

$$\sup_{t \in (0, T]} |(u - u^h)(\cdot, t)|_{1, \omega} \leq C (h/p)^S (\|\partial_t u\|_{L^2(H^S)} + \|u\|_{L^\infty(H^{S+1})}). \quad (5.43)$$

Theorem 5.5 Similarly, our error analysis of the discrete-time methods defined in Section 5.1.2 that use finite differences in time and FVE in space (as in fully-implicit “time-marching” procedures) yields the following error estimate:

$$\begin{aligned} \max_{1 \leq n \leq NT} |(u - u^h)(\cdot, t^n)|_{1, \omega} \\ \leq C \left(\Delta t^R \|\partial_t^{R+1} u\|_{L^2(L^2)} + \right. \\ \left. (h/p)^S (\|\partial_t u\|_{L^2(H^S)} + \|u\|_{L^\infty(H^{S+1})}) \right), \end{aligned} \quad (5.44)$$

where $S \in (1/2, \rho]$ and $R = 1$ or 2 depending on whether backward Euler ($R = 1$) or Crank-Nicolson ($R = 2$) time differencing is used.

In analogy to our elliptic analysis of Chapters 3 and 4, we note that the FVE parabolic error estimates (5.43) and (3.3) match or exceed the corresponding FE error estimates [55].

This concludes our presentation of FVE analysis. In the last two sections, the main results of the thesis are summarized and future work based on this thesis is outlined.

5.4 Summary

An *a priori* FVE error analysis for elliptic partial differential equations with smooth coefficients has been presented: $H^1(\Omega)$ -equivalent optimal-order and superconvergence results have been obtained in the $H^1(\omega)$ seminorm. These results can be extended to problems with discontinuous and nonlinear coefficients: the modifications for discontinuous coefficients are analogous to those of Samarskii *et al.* [46] for FV; the modifications for nonlinear coefficients are analogous to those of Douglas [12] for FE.

For the purpose of analysis, we have viewed FVE as a locally conservative Petrov-Galerkin FE and performed an FE-style error analysis. Building upon the foundation laid by Cai, Mandel, and McCormick [9] for steady diffusion equations discretized by linear finite elements in two spatial dimensions, we have extended and generalized FVE analysis to cover steady advection-reaction-diffusion equations discretized by C^0 polynomial finite elements of arbitrary order in three spatial dimensions. In addition, we have considered h , p , and h - p version FVE where increased accuracy is obtained by refining the computational mesh parametrized by h and/or by increasing the order p of the polynomial trial space.

In Sections 1.1 and 1.2 of Chapter 1, we outlined the origin of fluid dynamics and its basis on local conservation laws. Also, we motivated the need for FVE which preserves local conservation on the discrete level. In addition, we saw that FVE systematically discretizes the unknown solution, known source term data and boundary conditions on domains with complex geometries via the FE triangulation \mathcal{T}^h , the FVE triangulation $\mathcal{T}^{h/p}$, and the

FVE volumization $\mathcal{V}^{h/p}$, as discussed in Sections 2.1 and 2.2.

Motivated by the success of h , p , and h - p version FE in the structural mechanics literature (see [2, 51]), we defined h , p , and h - p version FVE for fluid dynamics problems in Section 2.3. Given the documented [39] success of Multi-level solution algorithms for h -version FVE vs. FE discretizations and the documented [25, 35, 37, 42] success of similar Multi-level and Domain Decomposition solution algorithms for p and h - p version FE discretizations, we expect similar implementation efficiency successes for p and h - p version FVE discretizations.

Furthermore, we saw that the computational set-up costs for FVE were less expensive than the costs for FE in Section 2.4. That is, if d is the spatial dimension of the problem domain, the quadrature rules used to evaluate FVE integrals and set-up a system of linear equations were seen to be $O(2^d)$ less expensive than corresponding FE quadrature rules: the quadrature operation count for reaction and source term integrals and is $O(p)^d$ for FVE and $O(2p)^d$ for FE. However, for reaction and source terms, FVE integrals are defined on volumes from $\mathcal{V}^{h/p}$ with diameters of $O(h/p)$ whereas FE integrals are defined on elements from \mathcal{T}^h with diameters of $O(h)$; therefore, the FVE quadratures are performed on regions that are $O(p)^d$ smaller than their FE counterparts. This feature makes the FVE equations more local in character and the FVE mass matrix arising from the reaction term more strongly diagonally dominant than its FE counterpart; from the standpoint of numerical linear algebra and implementation, this means that it should be easier to precondition and iteratively solve FVE systems of linear equations than their FE counterparts. In

addition, FVE diffusive and advective fluxes are defined as surface integrals which are simpler (i.e., of lower dimension) than the corresponding volume integrals for FE: the quadrature operation count is $O(p)^{d-1}$ for FVE and $O(2p)^d$ for FE—a increased savings of $O(p2^d)$.

Therefore, we expected FVE to outperform FE in standard measures of computational efficiency. Furthermore, from the standpoint of the *a priori* error analysis presented in the Chapters 3 and 4, we saw that the accuracy of FVE approximations to the exact solution of general elliptic equations either matched or surpassed FE asymptotic convergence rates for h , p , and h - p versions of the respective methods.

After introducing a Sobolev space infrastructure and a standard h , p , and h - p version FE analysis in Chapter 2, we detailed an h , p , and h - p version FVE analysis for diffusion equations in Chapter 3.

In Sections 3.1 and 3.2, we motivated and laid the groundwork for an FE-style analysis for FVE based on the formalisms introduced in the FE analysis presented in the previous chapter. After introducing discrete L^2 and H^1 Sobolev norms in Section 3.3, we demonstrated an upper bound for the FVE B -form in terms of our discrete H^1 semi-norm and an FVE diffusion functional in Section 3.4. After introducing the Bramble-Hilbert lemma in Section 3.5, we focused on an $H^1(\Omega)$ -equivalent h , p , and h - p version optimal-order approximation theory and an h -version superconvergent approximation theory for our FVE diffusion functional in Section 3.6. After establishing an optimal-order result in Theorem 3.1 for interpolation by C^0 piecewise polynomials of arbitrary

order on a regular triangulation and volumization, we focused on superconvergence for linear interpolation on a “symmetric” volumization in Theorem 3.2. That is, from the perspective of the FVE diffusion functional, the approximating power of a linear trial space on a symmetric volumization matches the power of a quadratic trial space on a regular volumization. At the core of the superconvergence theory for the diffusive flux functional is the 1-d fact that centered differencing is exact for quadratics; we invoke certain volume symmetries, discussed below, to extend this 1-d result to multi-dimensional settings.

As we saw in Sections 3.6.2 and 3.6.2.2, the definition of volume symmetry is dimensionally and geometry dependent. That is, if $d = 1$, volume symmetry is X -symmetry: the volume boundary γ_{ij} , a point in 1-d, is at the midpoint of each element edge X_{ij} ; in this simple case, the diffusive flux is represented by a centered difference quotient. If $d = 2$, volume symmetry for C^0 linears on triangles requires X and γ -symmetries: the volume edge γ_{ij} and the element edge X_{ij} are perpendicular bisectors of one another; however for C^0 bilinears on rectangles, volume symmetry requires only X -symmetry, since the quadratic component of a piecewise bilinear trial space obviated the γ_{ij} -symmetry requirement.

When $d = 3$, volume symmetry for C^0 linears on tetrahedra requires X -symmetry (the planar volume interface γ_{ij} is a perpendicular bisector for the element edge X_{ij}) and a “centroid” γ -symmetry: i.e., $X_{ij} - \gamma_{ij}$ and the point of intersection P_{ij} is the centroid for γ_{ij} ; furthermore, the geometry of γ_{ij} is such that the centroid quadrature rule is exact for linear functions— γ_{ij} is triangular,

rectangular, hexagonal, etc. (see [49] for details). For C^0 bilinears (i.e., $\Pi_{1,1}u \in \text{span}\{1, x, y\} \otimes \{1, z\}$) on triangular prisms formed in a tensor-product mesh, volume symmetry requires X -symmetry and a “2-d” γ -symmetry: the x and y cross-sections of γ_{ij} must possess the centroid symmetry. That is, linearity in x and y requires the 2-d γ_{ij} -symmetry in x and y , but bilinearity in z eliminates the 2-d γ_{ij} -symmetry in z . For C^0 trilinears on rectangular bricks formed by a tensor-product mesh, volume symmetry requires X -symmetry and a “1-d” γ_{ij} -symmetry: either one of the 1-d cross-sections of γ_{ij} must possess the centroid symmetry. That is, the quadratic components of a trilinear trial space obviate γ_{ij} -symmetry requirements as in 2-d; however, its cubic component invokes a reduced, 1-d γ_{ij} -symmetry.

Finally, using the “degenerate volume paradigm” of Section 3.6.2.2, we see that linear FVE on a symmetric volumization is identical to a quadratic FVE on a degenerate volumization; hence, superconvergence for linear FVE can be viewed as a special case of the optimal-order convergence for quadratic FVE.

When discussing superconvergence for linear trial spaces, it is important to emphasize that it is primarily a local phenomenon. Qualitatively, global superconvergence is the cumulative effect of local volume symmetries or uniformities in a given computational mesh. Quantitatively, local quadratic convergence leads to superlinear global convergence: the precise rate depends upon the prevalence of local volume symmetry or uniformity conditions on a global computational mesh. As a result, local quadratic convergence of a numerical solution (in regions of symmetry/uniformity on a non-uniform mesh)

“remediates” the global solution: superlinear, rather than linear, convergence is observed globally.

Using the optimal-order approximation results of Section 3.6.1, we derived (for $p \geq 2$) an “asymptotic” ellipticity result for the diffusion bilinear form in terms of our discrete FVE trial space $\mathcal{S}^p(\mathcal{T}^h)$ and our discrete H^1 semi-norm in Section 3.7. That is, for (h/p) sufficiently small, the ellipticity constant is independent of the computational mesh. This type of argument is unavoidable if we wish to maintain the FE formalism in our FVE error analysis for higher-order ($p \geq 2$) polynomial trial spaces, since our discrete H^1 semi-norm is defined as the restriction of the continuous H^1 semi-norm to piecewise linear polynomials on the FVE triangulation $\mathcal{T}^{h/p}$. Finally, in Section 3.8, we demonstrate an FE-style error analysis for h , p , and h - p version FVE that matches or exceeds the results of the standard h , p , and h - p FE analysis presented in Chapter 2.

In Chapter 4, we extended the scope of our analysis to general elliptic equations and detailed the analogous theoretical results for the FVE treatment of advection and reaction. After deriving upper bounds for the FVE advection and reaction bilinear forms in terms of our discrete norms and FVE advection and reaction functionals in Section 4.2, we focused our attention on developing an approximation theory for our advection and reaction functionals. In order to avoid potential conflicts with the superconvergence result for the FVE diffusion functional, we developed an h , p , and h - p version $L^2(\Omega)$ -equivalent optimal-order approximation theory for our advection and reaction functionals

in Section 4.3. Using these approximation results, we demonstrated “asymptotic” ellipticity results in terms of our discrete FVE trial space $\mathcal{S}^p(\mathcal{T}^h)$ and our discrete L^2 (and H^1) norm(s) for the reaction and advection bilinear forms in Section 4.4. That is, for (h/p) sufficiently small the ellipticity constants are independent of the computational mesh. Again this type of argument is necessary to maintain the FE formalism in our FVE analysis, since our discrete L^2 is defined to be the restriction of the continuous L^2 norm to piecewise constant functions on $\mathcal{V}^{h/p}$. In Section 4.5, we demonstrate an FE-style error analysis for steady reaction-diffusion and advection-reaction-diffusion equations for h , p , and h - p version FVE that matches or exceeds the corresponding FE results. And in the previous section, we concluded our presentation for elliptic equations with a discussion and analysis of the FVE elliptic projection operator to set the stage for an h , p , and h - p version FVE analysis for parabolic equations.

In the present chapter, the elliptic analysis presented here was extended to general parabolic equations in a straightforward manner by an FVE variant of the elliptic projection argument of Wheeler [55]. That is, an FE-style error analysis of a continuous-time method that uses ordinary differential equations in time and FVE in space (as in a “method of lines” procedure) yielded error estimates of the form

$$\sup_{t \in (0, T]} |(u - u^h)(\cdot, t)|_{1, \omega} \leq C (h/p)^S (\|\partial_t u\|_{L^2(H^S)} + \|u\|_{L^\infty(H^{S+1})}); \quad (5.45)$$

similarly, an FE-style error error analysis of a method that uses finite differences in time and FVE in space (as in a fully-implicit “time-marching” procedure)

yielded error estimates of the form

$$\begin{aligned}
& \max_{1 \leq n \leq N} |(u - u^h)(\cdot, t^n)|_{1,\omega} \\
& \leq C \left(\Delta t^R \|\partial_t^{R+1} u\|_{L^2(L^2)} + \right. \\
& \quad \left. (h/p)^S (\|\partial_t u\|_{L^2(H^S)} + \|u\|_{L^\infty(H^{S+1})}) \right), \quad (5.46)
\end{aligned}$$

where $S \in (1/2, \rho]$ and $R = 1$ or 2 depending on whether backward Euler ($R = 1$) or Crank-Nicolson ($R = 2$) time differencing is used.

5.5 Future Work

The work in this dissertation leads to the following key areas for future work (listed in order of relative importance and/or completeness):

- (1) basic theory for elliptic equations,
- (2) theory for advection–dominated parabolic equations and their solution by Eulerian-Lagrangian methods,
- (3) theory for porous media problems.

On the basic theory for elliptic equations, the key areas for development include:

- upwinding for advection,
- general boundary conditions,
- discrete L^2 analysis.

Preliminary work for upwinding with a linear ($p = 1$) trial space has been completed for locally one-dimensional or “operator-splitting” schemes and for a multi-dimensional or “full-operator” scheme similar to the algorithm developed, but not analyzed, by Otto [41] for FV. General boundary conditions pose no significant obstacles: non-homogeneous Dirichlet boundary conditions are easily included; total flux boundary conditions require a more detailed treatment of volume construction and symmetry on the domain boundary than what was stated here; also, preservation of a discrete compatibility condition is a crucial concern. A discrete L^2 analysis with a linear trial space has been completed that is analogous to the theory of Lazarov *et al.* [33] for FV which is based on a discrete variant of the Aubin-Nitsche duality lifting argument. One important difference between the two works must be noted, the success of [33]

for FV is based on uniform Cartesian-product meshes and the tone of the paper seems to indicate that similar success for general meshes is likely; however, the corresponding FVE analysis posed on general meshes proves that discrete L^2 success is limited to the uniform Cartesian-product mesh case—supporting numerical evidence indicates that this limitation holds for FV as well.

On the theory for advection-dominated parabolic equations and their solution by Eulerian-Lagrangian methods, the key areas for development include:

- MMOC/ELLAM,
- control volume ELLAM,
- flux-based ELLAM.

Following the theory developed here for parabolic equations, a theory for an FVE variant of the FE “modified method of characteristics” (MMOC) developed and analyzed by Douglas and Russell [15] has been completed under the assumption of a semi-analytic backtracking rule based on a Raviart-Thomas linear trial space representation (see Ciarlet [11] under “mixed finite element methods” for details) of the velocity field. The corresponding theory for an FVE variant of the more modern and systematic FE “Eulerian-Lagrangian localized adjoint method” (ELLAM) developed and analyzed by Wang, Ewing, and Russell [53] is in the works that includes a systematic treatment of boundary conditions. This theory will lay the ground work for two specific algorithms: the “control volume ELLAM” of Healy and Russell [22] and the “flux-based ELLAM” of Trujillo [52].

On the theory for porous media problems, the key areas for development include:

- FVE MMOC/ELLAM for transport equations,
- mixed FVE for flow equations.

The ultimate goal of FVE theory would be the development of an FVE variant of the FE theory of Russell [44] for coupled transport and flow problems in porous media. Analysis of the FVE methods for the transport equation directly builds on the analysis presented here. However, significant modifications must be made to analyze the “mixed FVE” (an FVE variant of the “mixed finite element methods” cited above) which has been developed, but not analyzed, by Jones [29].

As can be discerned from this brief discussion, significant advances beyond this dissertation await future development: those advances already completed or in progress will be presented elsewhere in the near future; however, advances in porous media problems may not be fully realized until the end of the next decade.

REFERENCES

- [1] R. A. Adams, *Sobolev Spaces*, Academic Press, New York, 1975.
- [2] I. Babuška and J. Osborn, *Eigenvalue Problems* in the *Handbook of Numerical Analysis: Vol. II, Finite Element Methods (Part 1)*, P. G. Ciarlet and J. L. Lions, ed., Elsevier (North-Holland), Amsterdam, 1991.
- [3] I. Babuška and M. Suri, “The optimal convergence rate of the p -version of the finite element method,” *SIAM J. Numer. Anal.*, **24** (1987), pp. 750–776.
- [4] B. R. Baliga and S. V. Patankar, “A new finite element formulation for convection-diffusion problems,” *Numer. Heat Transfer*, **3** (1980), pp. 393–409.
- [5] R. E. Bank and D. J. Rose, “Some error estimates for the box method,” *SIAM J. Numer. Anal.*, **24** (1987), pp. 777–787.
- [6] G. K. Batchelor, *An Introduction to Fluid Dynamics*, Cambridge Univ. Press, U.K., 1967.
- [7] J. H. Bramble and S. R. Hilbert, “Estimation of linear functionals with application to Fourier transforms and spline interpolation,” *SIAM J. Numer. Anal.*, **7** (1970), pp. 112–124.
- [8] Z. Cai and S. F. McCormick, “On the accuracy of the finite volume element method for diffusion equations on composite grids,” *SIAM J. Numer. Anal.*, **27** (1990), pp. 636–655.
- [9] Z. Cai, J. Mandel, and S. F. McCormick, “The finite volume element method for diffusion equations on general triangulations,” *SIAM J. Numer. Anal.*, **28** (1991), pp. 392–402.
- [10] Z. Cai, “On the finite volume element method,” *Numer. Math.*, **58** (1991), pp. 713–735.

- [11] P. G. Ciarlet, *Basic Error Estimates for Elliptic Problems* in the *Handbook of Numerical Analysis: Vol. II, Finite Element Methods (Part 1)*, P. G. Ciarlet and J. L. Lions, ed., Elsevier (North-Holland), Amsterdam, 1991.
- [12] J. Douglas, Jr., *Lectures on finite element methods for elliptic and parabolic problems*, Univ. of Chicago, 1973-74.
- [13] J. Douglas, Jr. and T. Dupont, "Galerkin methods for parabolic equations," *SIAM J. Numer. Anal.*, **7**, No. 4 (1970), pp. 575–626.
- [14] J. Douglas, Jr. and T. Dupont, "Galerkin methods for parabolic equations with nonlinear boundary conditions," *Numer. Math.*, **20** (1973), pp. 213–237.
- [15] J. Douglas, Jr. and T. F. Russell, "Numerical methods for convection-dominated diffusion problems based on combining the method of characteristics with finite element or finite difference procedures," *SIAM J. Numer. Anal.*, **19** (1982), pp. 871–885.
- [16] T. Dupont and R. Scott, "Polynomial approximation of functions in Sobolev spaces," *Math. Comp.*, **24**, No. 150 (1980), pp. 441–463.
- [17] R. E. Ewing, R. D. Lazarov, and P. S. Vassilevski, "Finite difference schemes on grids with local refinement in time and space for parabolic problems: I. derivation, stability, and error analysis," *Computing*, **45**, No. 3 (1990), pp. 193–215.
- [18] R. E. Ewing, R. D. Lazarov, and P. S. Vassilevski, "Local refinement techniques for elliptic problems on cell-centered grids: I. error analysis," *Math. Comp.*, **56** (1991), pp. 437–461.
- [19] P. Grisvard, *Elliptic Problems in Non-smooth Domains*, Pitman, Boston, 1985.
- [20] R. B. Guenther and J. W. Lee, *Partial Differential Equations of Mathematical Physics and Integral Equations*, Prentice-Hall, New Jersey, 1988.
- [21] W. Hackbusch, "On first and second order box schemes," *Computing*, **41** (1989), pp. 277–296.
- [22] R. W. Healy and T. F. Russell, "A conservative finite-volume Eulerian-Lagrangian localized adjoint method for solution of the advection-dispersion equation," *Water Resour. Res.* **29** (1993), pp. 2399–2413.

- [23] B. Heinrich, *Finite Difference Methods on Irregular Networks: generalized approach to second order elliptic problems*, International Series in Numerical Mathematics, **82**, Birkhäuser, Stuttgart, 1987.
- [24] R. Herbin, “An error estimate for a finite volume scheme for a diffusion-convection problem on a triangular mesh,” *Numer. Meth. PDE*, **11** (1995), pp. 165–173.
- [25] N. Hu and I. N. Katz, “Multi- p methods: iterative algorithms for the p -version of the finite element analysis,” *SIAM J. Sci. Comput.*, **16** (1995), pp. 1308–1332.
- [26] T. J. R. Hughes, ed., *Finite Element Methods for Convection Dominated Flows*, AMD **34**, ASME, 1979.
- [27] S. Jensen and M. Suri, “On the L^2 error of the p -version of the finite element method over polygonal domains,” *Comp. Meth. Appl. Mech. Engrg.*, **97** (1992), pp. 233–243.
- [28] C. Johnson, *Numerical Solution of Partial Differential Equations by the Finite Element Method*, Cambridge Univ. Press, 1987.
- [29] J. Jones, “A mixed finite volume element method for accurate computation of fluid velocities in porous media,” Ph. D. Thesis, University of Colorado, Denver, CO, 1995.
- [30] S. Kesavan, *Topics in Functional Analysis and Applications*, Wiley Eastern Limited, New Delhi, India, 1989.
- [31] H. O. Kreiss, T. A. Manteuffel, B. Swartz, B. Wendroff, and A. B. White, “Supraconvergent schemes on irregular grids,” *Math. Comp.*, **47** (1986), pp. 537–554.
- [32] R. D. Lazarov, I. D. Mishev, and P. S. Vassilevski, “Finite volume methods with local refinement for convection-diffusion problems,” *Computing*, **53** (1994), pp. 33–57.
- [33] R. D. Lazarov, I. D. Mishev, and P. S. Vassilevski, “Finite volume methods for convection-diffusion problems,” *SIAM J. Numer. Anal.*, **33** (1996), pp. 31–55.
- [34] J. Mandel, “Iterative solver for p -version finite elements in three dimensions,” *Comp. Meth. Appl. Mech. Engrg.*, **116** (1994), pp. 411–425.

- [35] T. A. Manteuffel and A. B. White, “The numerical solution of second-order boundary value problems on nonuniform meshes,” *Math. Comp.*, **47** (1986), pp. 511–535.
- [36] A. R. Mitchell and D. F. Griffiths, *The Finite Difference Method in Partial Differential Equations*, Wiley, New York, 1980.
- [37] W. R. Mitchell, “Optimal multilevel iterative methods for adaptive grids,” *SIAM J. Sci. Stat. Comput.*, **13** (1992), pp. 146–167.
- [38] K. W. Morton and E. Süli, “Finite volume methods and their analysis,” *IMA J. Numer. Anal.*, **11** (1991), pp. 241–260.
- [39] S. F. McCormick, *Multilevel Adaptive Methods for Partial Differential Equations*, Chapter 2: The Finite Volume Element Method, *Frontiers in Applied Mathematics*, Vol. 6, SIAM, Philadelphia, 1989.
- [40] J. T. Oden, *Finite Elements: An Introduction in the Handbook of Numerical Analysis: Vol. II, Finite Element Methods (Part 1)*, P. G. Ciarlet and J. L. Lions, ed., Elsevier (North-Holland), Amsterdam, 1991.
- [41] J. S. Otto, “Multilevel methods for the solution of advection–dominated elliptic problems on composite grids,” Ph. D. Thesis, University of Colorado, Denver, CO, 1992.
- [42] L. F. Pavarino, “Additive Schwarz methods for the p -version finite element method,” *Numer. Math.*, **66** (1994), pp. 493–515.
- [43] P. J. Roache, *Computational Fluid Dynamics*, Hermosa, Albuquerque, 1976.
- [44] T. F. Russell, “Time-stepping along characteristics with incomplete iteration for a Galerkin approximation of miscible displacement in porous media,” *SIAM J. Numer. Anal.*, **22** (1985), pp. 970–1013.
- [45] T. F. Russell and R. V. Trujillo, “The finite volume element method for elliptic and parabolic equations,” in *Numerical Analysis: A. R. Mitchell 75th Birthday Volume*. G. A. Watson and D. F. Griffiths, eds., World Scientific Publishing, 1996.
- [46] A. A. Samarskii, R. D. Lazarov, and V. L. Makarov, *Finite Difference Schemes for Differential Equations with Weak Solutions*, Vysshaya Shkola Publishers, Moscow, 1987. (In Russian.)

- [47] E. G. Sewell, "Automatic generation of triangulations for piecewise polynomial approximation," Ph. D. Thesis, Purdue University, West Lafayette, IN, 1972.
- [48] M. Shashkov and S. Steinberg, *Conservative Finite-Difference Methods on General Grids*, CRC Press, Boca Raton, 1996.
- [49] A. H. Stroud, *Approximate Calculation of Multiple Integrals*, Prentice-Hall, New Jersey, 1971.
- [50] E. Süli, "Convergence of finite volume schemes for Poisson's equation on nonuniform meshes," *SIAM J. Numer. Anal.*, **28** (1991), pp. 1419-1430.
- [51] B. A. Szabo and I. Babuška, *Finite Element Analysis*, J. Wiley & Sons, New York, 1991.
- [52] R. V. Trujillo, "Flux-based Eulerian-Lagrangian methods for transient advection equations," UCD/CCM Report No. 20, 1994.
- [53] H. Wang, R. E. Ewing, and T. F. Russell, "Eulerian-Lagrangian localized adjoint methods for convection-diffusion equations and their convergence analysis," *IMA J. Numer. Anal.*, **15** (1995), pp. 405-459.
- [54] A. Weiser and M. F. Wheeler, "On convergence of block-centered finite differences for elliptic problems," *SIAM J. Numer. Anal.*, **25** (1988), pp. 351-357.
- [55] M. F. Wheeler, "*A priori* L^2 error estimates for Galerkin approximations to parabolic partial differential equations," *SIAM J. Numer. Anal.*, **10** (1973), pp. 723-759.